

Fri May 9 14:20:38 2003

FEATURES		Location/Qualifiers	
source	1..4321	/organism="synthetic construct"	
BASE COUNT	966 a 1281 c 1168 g 906 t	/db_xref="taxon:32630"	
ORIGIN			
Query Match	82.4%;	Score 4188.2;	DB 6; Length 4321;
Best Local Similarity	99.1%;	Pred. No. 0;	
Matches 4297;	Conservative 0;	Mismatches 16;	Indels 24; Gaps 8;
Qy	754	CGAGATCCCGGAGCAGCTTCTGGAGAGGGGAACGGTCCGGAGCAAGCCAGAGG	813
Db	1	CGAGATCCCGGAGCAGCTTCTGGAGAGCGGG-ACGGTCCGGAGAGCAAGCCACAGG	59
Qy	814	CAGAGGAGGCGACAGAGGAGGAAAGGCGCCNAGCTAGCGCTCCAGTGTGTACAGNAGC	873
Db	60	CAGAGGAGGCGACAGAGGAGGAAAGGG-CCGAGCTAGCGCTCCAGTGTGTACAGGAGC	118
Qy	874	CGAA-GGACGCACACGCCAGCCCGGCTCCAGGACAGCAGCAGCAGCCTTGCA-	931
Db	119	CGAAGGAGCAGCAGCAGCAGCCAGCCAGCGGCTCCAGGAGACAGCAAGCCTCTTGCA	178
Qy	932	-----NCGGTTGCAAGCGCGCCCGGAGCTGCCCTTCTCTTCTGCTGAAAGTTTTTAAA	986
Db	179	CGCGGCGGCTTCGAAGCGCGCGCGCGAGCTGCCCTTCTCTTCTGCTGAAAGTTTTTAAA	238
Qy	987	AGCTGCTAAAGACTCGGAGGAGCAAGAAAGTGCCTGGTAGACTAGCGGTGCTTTG	1046
Db	239	AGCTGCTAAAGACTCGGAGGAGCAAGAAAGTGCCTGGTAGACTAGCGGTGCTTTG	298
Qy	1047	TCCTCTCTCTTCACCGCGCTCCCGCCAGCTGCTCCCGCCCTCCCGCCCTCTCTTC	1106
Db	299	TCCTCTCTCTTCACCGCGCTCCCGCCAGCTGCTCCCGCCCTCCCGCCCTCTCTTC	358
Qy	1107	TCCCGCAGCTGCTCAGTGGGTACTCTCAGCCAAACCCCGCTCACCACCTTCTCCGCC	1166
Db	359	TCCCGCAGCTGCTCAGTGGGTACTCTCAGCCAAACCCCGCTCACCACCTTCTCCGCC	418
Qy	1167	CGCGCCCGCCCGCGCTGCGGCCAGCGTNCAGCNCGGAGTTGCGAGAGGTAAGTCC	1226
Db	419	CGCGCCCGCCCGCGCTGCGGCCAGCGTGCAGCCCGAGTTGCGAGAGGTAAGTCC	478
Qy	1227	CTTTGGCTCGAGCGGGGAGNCTAGCTGCACATTGCAAGAGAGGCTCTTAGAG-	1285
Db	479	CTTTGGCTCGAGCGGGGAG-CTAGCTGCACATTGCAAGAGAGGCTCTTAGAGCCAGG	537
Qy	1286	CGACTGGGAGCGGCTTCAGACTCGAGCCAGCCNCGCTGGTTAGGCTGCACGCGGAG	1345
Db	538	CGACTGGGAGCGGCTTCAGACTCGAGCCAGCCNCGCTGGTTAGGTTCCGCGGAG	597
Qy	1346	AGAACCTCTGTTTCCCGCCACTCTCTCCACCTCTCTCTGCTTCCCGCCCGCCAGGTG	1405
Db	598	AGAACCTCTGTTTCCCGCCACTCTCTCTCCACCTCTCTCTGCTTCCCGCCCGCCAGGTG	657
Qy	1406	CGGAGCAGAGATCAAAAGATGAAAGGAGTCAGTCTTCAGTAGCCAAAAGCAAAAC	1465
Db	658	CGGAG-CAGAGATCAAAAGATGAAAGGAGTCAGTCTTCAGTAGCCAAAAGCAAAAC	716
Qy	1466	AAACAAAACAAAAGCCGCAATATAAGAAAAGATATAACTCAGTCTTATTTCGAC	1525
Db	717	AAACAAAACAAAAGCCGCAATATAAGAAAAGATATAACTCAGTCTTATTTCGAC	776
Qy	1526	CTACTTCAGTGGACACTGAATTTGGAAGTGGAGGATTTGTTTCTTTTAAAGATCT	1585
Db	777	CTACTTCAGTGGACACTGAATTTGGAAGTGGAGGATTTGTTTCTTTTAAAGATCT	836
Qy	1586	GGGATCTTTTGAATCTACCTTCAAGTATTAGAGACAGACTGTGAGCCTTAGCAGGCA	1645
Db	837	GGGATCTTTTGAATCTACCTTCAAGTATTAGAGACAGACTGTGAGCCTTAGCAGGCA	896
Qy	1646	GATCTTGTCACCGCTGCTCTCTCTGACGAGACTTTGAGGCTGTACAGCGCTTTT	1705

897	GATCTTGTCACCGCTGTGTCTTCTTCTGCACGAGACTTTTGGAGCTGTCTAGACGCGCTTTT	956.
1706	GCSTGGTTGCTCCCGCAAGTTTCTTCTCTGGAGCTTCCCGCAGSTGGGCAGCTAGCTGCG	1765
957	CGTGTTGTTCTCCCGCAAGTTTCTTCTCTGTGAGCTTCCCGCAGGTGGCAGCTAGCTGC	1016
1766	AGCGACTACCGCATCATCAGAGCCTGTGTGAACCTTCTTGTAGCAAGAGAAGGGAGCGGG	1825
1017	AGCGACTACCGCATCATCAGAGCCTGTGTGAACCTTCTTGTAGCAAGAGAAGGGAGCGGG	1076
1826	GTAGGGAAGTAGTGTGAAGATTTCAGCCAAAGCTCAAGGATGGAAGTGCAGTTTAGGGCTGG	1885
1077	GTAGGGAAGTAGTGTGAAGATTTCAGCCAAAGCTCAAGGATGGAAGTGCAGTTTAGGGCTGG	1136
1886	GAAGGGTCTACCTTCGCGCCGCTCCAGAGCTTACAGAGCTTCCAGAAATCTCTTCC	1945
1137	GAAGGGTCTACCTTCGCGCCGCTCCAGAGCTTACCGAGAGCTTTCCAGAAATCTGTTC	1196
1946	ACAGCGTTCGGCAAGTGATCCAGAAACCCGGGCCACAGGCACCCAGAGGCGCGAGCGCAG	2005
1197	ACAGCGTTCGGCAAGTGATCCAGAAACCCGGGCCACAGGCACCCAGAGGCGCGAGCGCAG	1256
2006	CACCTCCCGCGCCAGTTTGTCTGTCTGTGACGAGCAGCAGCAGCAGCAGCAGCAGCAGC	2065
1257	CACCTCCCGCGCCAGTTTGTCTGTCTGTGACGAGCAGCAGCAGCAGCAGCAGCAGCAGC	1304
2066	AGC	2125
1305	AGC	1364
2126	AGC	2185
1365	AGC	1424
2186	GCTACCTGTGCTTGGATGAGGAACAGCAACCTTTCACAGCCGCGAGCTTCGAGGTGCC	2245
1425	GCTACCTGTGCTTGGATGAGGAACAGCAACCTTTCACAGCCGCGAGCTTCGAGGTGCC	1484
2246	ACCCGAGAGAGGTTGCGTCCAGAGCCTTCGAGCCGCGAGGATGACTCAGCTGCCCGCTT	2305
1485	ACCCGAGAGAGGTTGCGTCCAGAGCCTTCGAGCCGCGAGGATGACTCAGCTGCCCGCTT	1544
2306	CGCAGCAGCTTCCGAGCAGCCTTCGAGCCGCGAGGATGACTCAGCTGCCCGCTT	2365
1545	CGCAGCAGCTTCCGAGCAGCCTTCGAGCCGCGAGGATGACTCAGCTGCCCGCTT	1604
2366	TGCTGGGCGCCACTTTCGCGGCTTAAAGCAGCTGCTTCCGCTCAGCTTAAAGACATCTGA	2425
1605	TGCTGGGCGCCACTTTCGCGGCTTAAAGCAGCTGCTTCCGCTCAGCTTAAAGACATCTGA	1664
2426	CGCAGGCGCAGCAGCATGCAACTCTTTCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGC	2485
1665	CGCAGGCGCAGCAGCATGCAACTCTTTCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGC	1724
2486	GCAGCAGGGGAGAGCAGGAGGAGGCTTCGCGGCTTCCACTTCCCAAGGACATTAATCT	2545
1725	GCAGCAGGGGAGAGCAGGAGGAGGCTTCGCGGCTTCCACTTCCCAAGGACATTAATCT	1784
2546	TAGGGGCACTTCACCATTTCTGACAAACCCCAAGGAGTTGTGTAGGAGCAGTGTCCGTT	2605
1795	TAGGGGCACTTCACCATTTCTGACAAACCCCAAGGAGTTGTGTAGGAGCAGTGTCCGTT	1844
2606	CCATGGGCGCTGGTGTGGAGCGTTTGGAGCTTCCAGGAGGAGCAGTTCGCGGGG	2665
1845	CCATGGGCGCTGGTGTGGAGCGTTTGGAGCTTCCAGGAGGAGCAGTTCGCGGGG	1904
2666	ATTGCATGTAGCCCGCCACTTTTGGGAGTTCCACCGCTGTGCGTCCCGCTTGTGCC	2725
1905	ATTGCATGTAGCCCGCCACTTTTGGGAGTTCCACCGCTGTGCGTCCCGCTTGTGCC	1964
2726	CATTGGCGAATGCAAGGTTCTCTGTAGACGACACGCGAGGCAAGC	

Db	897	GATCTTGTCACCGGTGTGTCTTCTTCTGACGAGACTTTGAGGCTGTACAGCGCTTTT 956	
Qy	1706	GGGTGTTGCTCCCGCAAGTTTCTTCTGAGCTTCCCGCAGGTGGGAGCTAGCTGTC 1765	
Db	957	GGGTGTTGCTCCCGCAAGTTTCTTCTGAGCTTCCCGCAGGTGGGAGCTAGCTGTC 1016	
Qy	1766	AGCGACTACCGCATCATCACAGCCTTGTAACTCTTGTGACAAAGAGAGGGAGCGGG 1825	
Db	1017	AGCGACTACCGCATCATCACAGCCTTGTAACTCTTGTGACAAAGAGAGGGAGCGGG 1076	
Qy	1826	GTAAAGGAAGTAGTGTGAAGATTTCAGCCAAAGTCAAGATGAAGTGCAGTTAGGGCTGG 1885	
Db	1077	GTAAAGGAAGTAGTGTGAAGATTTCAGCCAAAGTCAAGATGAAGTGCAGTTAGGGCTGG 1136	
Qy	1886	GAAGGCTTACCTTCGGCCCGCTCAAGACCTACCGAGGAGCTTTCAGAAATCTGTTC 1945	
Db	1137	GAAGGCTTACCTTCGGCCCGCTCAAGACCTACCGAGGAGCTTTCAGAAATCTGTTC 1196	
Qy	1946	AGAGCGTGGCGAAGTGTCCAGAACCCCGGCCCCAGCACCAGAGCGCGAGCGCAG 2005	
Db	1197	AGAGCGTGGCGAAGTGTCCAGAACCCCGGCCCCAGCACCAGAGCGCGAGCGCAG 1256	
Qy	2006	CACCTCCCGCGCCAGTTTGTCTGCTCTGACGACGACGACGACGACGACGACGACG 2065	
Db	1257	CACCTCCCGCGCCAGTTTGTCTGCT-----GCAGCAGCAGCAGCAGCAGC 1304	
Qy	2066	AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2125	
Db	1305	AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1364	
Qy	2126	AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2185	
Db	1365	AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1424	
Qy	2186	GCTACCTGGTCTTGATGAGAAACAGCACTTCACAGCGGAGTGCAGCGCTGGAGTGC 2245	
Db	1425	GCTACCTGGTCTTGATGAGAAACAGCACTTCACAGCGGAGTGCAGCGCTGGAGTGC 1484	
Qy	2246	ACCCGAGAGAGGTTGCTCCAGAGCTGAGCGCGCTGGCGCCAGCAGAGGGCTGC 2305	
Db	1485	ACCCGAGAGAGGTTGCTCCAGAGCTGAGCGCGCTGGCGCCAGCAGAGGGCTGC 1544	
Qy	2306	CGCAGCAGCTGCCAGCAGCTCCCGAGCAGGATGACTCAGCTGCCCTCCATCCACGTTCTCC 2365	
Db	1545	CGCAGCAGCTGCCAGCAGCTCCCGAGCAGGATGACTCAGCTGCCCTCCATCCACGTTCTCC 1604	
Qy	2366	TGCTGGGCCCCACTTTCGCCGCTTAAAGAGCTGCTCCGCTGACCTTAAAGACATCTGA 2425	
Db	1605	TGCTGGGCCCCACTTTCGCCGCTTAAAGAGCTGCTCCGCTGACCTTAAAGACATCTGA 1664	
Qy	2426	GGGAGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2485	
Db	1665	GGGAGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1724	
Qy	2486	GCAGCAGCGGGAGAGCGAGGAGGCGCTCGGGGCTCCACCTTCTTCCAAAGCAAAATTA 2545	
Db	1725	GCAGCAGCGGGAGAGCGAGGAGGCGCTCGGGGCTCCACCTTCTTCCAAAGCAAAATTA 1784	
Qy	2546	TAGGGGGCACTTCGACCACTTCTGACAAAGCGCAAGAGTGTGTGAAGCAGTGTGCTGT 2605	
Db	1785	TAGGGGGCACTTCGACCACTTCTGACAAAGCGCAAGAGTGTGTGAAGCAGTGTGCTGT 1844	
Qy	2606	CCATGGCGCTGGGTGTGAGGAGCTTGGAGCATCTGAGTCCAGGGGACAGCTTCGGGGG 2665	
Db	1845	CCATGGCGCTGGGTGTGAGGAGCTTGGAGCATCTGAGTCCAGGGGACAGCTTCGGGGG 1904	
Qy	2666	ATTGCAATGTACGCCCCACTTTTGGGAGTTCACCCGCTGTGCGTCCCACTCTTGTGCCC 2725	
Db	1905	ATTGCAATGTACGCCCCACTTTTGGGAGTTCACCCGCTGTGCGTCCCACTCTTGTGCCC 1964	
Qy	2726	CATTGGCCGAATCAAGAGTCTCTGCTAGACAGCAGCGCAGGAGCAGTCAAGATA 2785	
Db	1965	CATTGGCCGAATCAAGAGTCTCTGCTAGACAGCAGCGCAGGAGCAGTCAAGATA 2024	

QY	2786	CTGCTGAGTATCCCTTTCAAGGAGGATTACACAAAGGGCTAGAAAGCGAGAGCCCTAG	2845
Db	2025	CTGCTGAGTATTCCTTTCAAGGAGGATTACACAAAGGGCTAGAAAGCGAGAGCCCTAG	2084
QY	2846	GCTGCTCTGCGAGCGCTCCACAGGAGACTCCGGGACACTTGAATGCGCTCTACCCCTGT	2905
Db	2085	GCTGCTCTGCGAGCGCTCCACAGGAGACTCCGGGACACTTGAATGCGCTCTACCCCTGT	2144
QY	2906	CTCTCTACAAGTCCGGAGCACTGCACGAGGACGTCGCTACAGAGTCCGACACTACTACA	2965
Db	2145	CTCTCTACAAGTCCGGAGCACTGCACGAGGACGTCGCTACAGAGTCCGACACTACTACA	2204
QY	2966	ACTTTTCCACTGCTCTGCGCGGACCGCCGCCCTCCGCGGCTCCCATCCCCACGCTC	3025
Db	2205	ACTTTTCCACTGCTCTGCGCGGACCGCCGCCCTCCGCGGCTCCCATCCCCACGCTC	2264
QY	3026	GCATCAAGCTGGAGAACCCGCTGCACTACGACAGCGCTGGCGGCTGCGCGGCGCAGT	3085
Db	2265	GCATCAAGCTGGAGAACCCGCTGCACTACGACAGCGCTGGCGGCTGCGCGGCGCAGT	2324
QY	3086	GCGCTATTGGGACCTGGCGAGCCTGCATGGCGGGGTGCAGCGGGACCCGGTTCTGGGT	3145
Db	2325	GCGCTATTGGGACCTGGCGAGCCTGCATGGCGGGGTGCAGCGGGACCCGGTTCTGGGT	2384
QY	3145	CACCTCTAGCCGCGCTTCCTCATCTGCGCACACTCTCTTCACAGCGAAGAGGCCAGT	3205
Db	2385	CACCTCTAGCCGCGCTTCCTCATCTGCGCACACTCTCTTCACAGCGAAGAGGCCAGT	2444
QY	3206	TGATATGACCTGTGTGTGTGGGTGTGTGGCGGGGTGTGGCGGGCGGGCGGGCGGGG	3265
Db	2445	TGATATGACCTGTGTGTGTGGGTGTGTGGCGGGGTGTGGCGGGCGGGCGGGCGGGG	2504
QY	3266	GCGGGGGGGCGGGCGGGCGGGGAGGGGAGTGTAGCCCTCTAGCGCTACGCTACATCTC	3325
Db	2505	GCGGGGGGGCGGGCGGGCGGGGAGGGGAGTGTAGCCCTCTAGCGCTACGCTACATCTC	2564
QY	3326	GGCCCCCTCAGGGGCTGGCGGCCAGGAAAGGCACTTACCGCACTGTGTGTGTGCTACC	3385
Db	2565	GGCCCCCTCAGGGGCTGGCGGCCAGGAAAGGCACTTACCGCACTGTGTGTGTGCTACC	2624
QY	3386	CTGCGGCGATGTTGAGCAGAGTGCCTATCCACGTGCCACTTGTGTGTGTGTGTGTGTG	3445
Db	2625	CTGCGGCGATGTTGAGCAGAGTGCCTATCCACGTGCCACTTGTGTGTGTGTGTGTGTG	2684
QY	3446	GCCCTGGAATGATAGCTACTCCGGACCTTACGGGGACATTCGCTTTGGAGACTGCCAGG	3505
Db	2685	GCCCTGGAATGATAGCTACTCCGGACCTTACGGGGACATTCGCTTTGGAGACTGCCAGG	2744
QY	3506	ACCATGTTTTGCCCATTGACTATTACTTTCACCCAGAGACCTTCCTGTATCTGTGGAG	3565
Db	2745	ACCATGTTTTGCCCATTGACTATTACTTTCACCCAGAGACCTTCCTGTATCTGTGGAG	2804
QY	3566	ATCAAGCTTCTGGGTGTCATATGGAGCTCTACATGTGGAAGCTGCAAGGCTCTCTTCA	3625
Db	2805	ATGAAGCTTCTGGGTGTCATATGGAGCTCTACATGTGGAAGCTGCAAGGCTCTCTTCA	2864
QY	3626	AAAGAGCCGCTCAAGGGAAACAGAAAGTACCTGTGCGCCAGCAGAGAAATGATTCGCTATTG	3685
Db	2865	AAAGAGCCGCTCAAGGGAAACAGAAAGTACCTGTGCGCCAGCAGAGAAATGATTCGCTATTG	2924
QY	3686	ATAAATTCCGAGGAAAAATTTGCTATTTGCTCTTCGAAATGTTATGAAGCAGGGA	2984
Db	2925	ATAAATTCCGAGGAAAAATTTGCTATTTGCTCTTCGAAATGTTATGAAGCAGGGA	3044
QY	3746	TGACTCTGGAGCCCGGAAGCTGAGAAACTTGGTAATCTGAACTACAGAGGAGGAG	3805
Db	2985	TGACTCTGGAGCCCGGAAGCTGAGAAACTTGGTAATCTGAACTACAGAGGAGGAG	3044
QY	3806	AGGCTTCCAGCACCAAGCCGCTGAGGAGAACACCAAGCTGACAGTGTACACAC	3865
Db	3045	AGGCTTCCAGCACCAAGCCGCTGAGGAGAACACCAAGCTGACAGTGTACACAC	3104

QY	3866	TTGAAGGCTATGAATGTACAGCCCACTTTTCTCAATGTCTCTGGAAGCCATTTAGCCAGGTG	3925
DB	3105	TTGAAGGCTATGAATGTGACGCCCATCTTCTGAATGTCTCTGGAAGCCATTTAGC	3164
QY	3926	TAGTGTGTCTGGACACGACAAACACCCAGCCGACTCTCTTGCAGCCCTTCTCTATGCC	3985
DB	3165	TAGTGTGTCTGGACACGACAAACACCCAGCCGACTCTCTTGCAGCCCTTCTCTATGCC	3224
QY	3986	TCAATGAACCTGGAGAGACAGCTGTGTACAGTGGTCAAGTGGCCNAGGCCCTTGCGTG	4045
DB	3225	TCAATGAACCTGGAGAGACAGCTGTGTACAGTGGTCAAGTGGCCNAGGCCCTTGCGTG	3284
QY	4046	GCTTCGGCAACTTACAGTGGAGCACCAGATGGCTGTCAATTCAGTACTCTGTGATGGGC	4105
DB	3285	GCTTCGGCAACTTACAGTGGAGCACCAGATGGCTGTCAATTCAGTACTCTGTGATGGGC	3344
QY	4106	TCATGGTCTTTGCCATGGCTGGCGATCCTTACCAATGTCAACTCCAGATGCTCTACT	4165
DB	3345	TCATGGTCTTTGCCATGGCTGGCGATCCTTACCAATGTCAACTCCAGATGCTCTACT	3404
QY	4166	TCGCCCTGTATCTGGTTTCAATGAGTACCGCATGCAAAAGTCCCGGATGTACAGCCAGT	4225
DB	3405	TCGCCCTGTATCTGGTTTCAATGAGTACCGCATGCAAAAGTCCCGGATGTACAGCCAGT	3464
QY	4226	GTGTCCGAATGAGCACACTCTCTCAAGAGTTTGGATGGCTCCAAATCACCCCCAGGAAT	4285
DB	3465	GTGTCCGAATGAGCACACTCTCTCAAGAGTTTGGATGGCTCCAAATCACCCCCAGGAAT	3524
QY	4286	TCCGTGTCATGAAGCAGCTGCTACTCTTCAAGAGTTTGGATGGCTCCAAATCACCCCCAGGAAT	4345
DB	3525	TCCGTGTCATGAAGCAGCTGCTACTCTTCAAGAGTTTGGATGGCTCCAAATCACCCCCAGGAAT	3584
QY	4346	AAAAATCTTTGATGAATCTCGAATGAACATACATCAAGGAACCTCGATCTATCATTTGCAT	4405
DB	3585	AAAAATCTTTGATGAATCTCGAATGAACATACATCAAGGAACCTCGATCTATCATTTGCAT	3644
QY	4406	GCAAGAAATAATCCACATCTCTGCTCAAGCGCTTCCACAGCTCACCAAGCTCCTGG	4465
DB	3645	GCAAGAAATAATCCACATCTCTGCTCAAGCGCTTCCACAGCTCACCAAGCTCCTGG	3704
QY	4466	ACTCCGTGACGCTATTGGGAGAGCTGCATCAGTTCATCTTTCACCTGCTTAATCAAGT	4525
DB	3705	ACTCCGTGACGCTATTGGGAGAGCTGCATCAGTTCATCTTTCACCTGCTTAATCAAGT	3764
QY	4526	CACACATGTGACGCTGGACTTTCGGAAATGATGCAGAGATCATCTCTGTGCAAGTGC	4585
DB	3765	CACACATGTGACGCTGGACTTTCGGAAATGATGCAGAGATCATCTCTGTGCAAGTGC	3824
QY	4586	CCAAGATCTCTTCTGCGAAAGTCAAGCCCATCTATTCCACACCAGTGAAGCATTTGGA	4645
DB	3825	CCAAGATCTCTTCTGCGAAAGTCAAGCCCATCTATTCCACACCAGTGAAGCATTTGGA	3884
QY	4646	ACCCATATTTCCACACCAGCTCATATGCCCCCTTTCCAGATGCTTCTGCTCTTATACCTC	4705
DB	3885	ACCCATATTTCCACACCAGCTCATATGCCCCCTTTCCAGATGCTTCTGCTCTTATACCTC	3944
QY	4706	TGCATCTACTCTCTCAGTGGCTTGGGAAATTTCTCTTCTCTCTCTCTCTCTCTCTCTCT	4765
DB	3945	TGCATCTACTCTCTCAGTGGCTTGGGAAATTTCTCTTCTCTCTCTCTCTCTCTCTCTCT	4004
QY	4766	ACATGTCTCTGAATTCATTTGCTGGGCTTTTTTTTTCTCTCTCTCTCTCTCTCTCTCTCT	4825
DB	4005	ACATGTCTCTGAATTCATTTGCTGGGCTTTTTTTTTCTCTCTCTCTCTCTCTCTCTCTCT	4064
QY	4826	TCTTCCCTCCCTATCTAACCCCTCCATGGACCTTCAGACTTTGCTTCCCATTTGGCTC	4885
DB	4065	TCTTCCCTCCCTATCTAACCCCTCCATGGACCTTCAGACTTTGCTTCCCATTTGGCTC	4124
QY	4886	CTATCTGTGTTTGAATGGTGTGTATGCGCTTTTAAATCTGTGATGATCCTCATATGGCCC	4945
DB	4125	CTATCTGTGTTTGAATGGTGTGTATGCGCTTTTAAATCTGTGATGATCCTCATATGGCCC	4184
QY	4946	AGTGTCAAGTGTGCTTTGTTTACAGCACTACTCTGTGCCAGCACAAACGTTTACTTTA	5005

Db	543	GTTAGGGCTGGGAAGGGTTACCTCTGGCGCGCTCAAGACTTACCGAGGAGCTTTCCA	602
QY	1935	GAATCTGTCCACAGCGTGGCGAAGTGATCCAGAACCCGGGCCCAAGCACCAGAGGC	1994
Db	603	GAATCTGTTCACAGCGTGGCGAAGTGATCCAGAACCCGGGCCCAAGCACCAGAGGC	662
QY	1995	CGGAGCGCAGCACCTCCCGGGCCCAAGTTTGTCTGTCTGCAGCAGCAGCAGCAGCA	2054
Db	663	CGGAGCGCAGCACCTCCCGGGCCCAAGTTTGTCTGTCTGCAGCAGCAGCAGCAGCA	701
QY	2055	GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	2114
Db	702	---GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	758
QY	2115	TAGCCCCAGGCAGCAGCAGCAGCAGCGGTGAGGATGTTCTCCCAAGCCCATCGTAG	2174
Db	759	TAGCCCCAGGCAGCAGCAGCAGCAGCGGTGAGGATGTTCTCCCAAGCCCATCGTAG	818
QY	2175	AGGCCCCACAGGCTACCTGGTCTGGATGAGGAACCAACCTTCACAGCCGACGTGGC	2234
Db	819	AGGCCCCACAGGCTACCTGGTCTGGATGAGGAACCAACCTTCACAGCCGACGTGGC	878
QY	2235	CCTGAGGTGCCACCCCGAGAGAGGTTCGCTCCAGAGCCTGGAGCCCGCTGGCCGCCAG	2294
Db	879	CCTGAGGTGCCACCCCGAGAGAGGTTCGCTCCAGAGCCTGGAGCCCGCTGGCCGCCAG	938
QY	2295	CAAGGGGTGGCGAGCAGCTGCCAGCACCTCCCGGACGAGATGACTCAGCTGCCCATC	2354
Db	939	CAAGGGGTGGCGAGCAGCTGCCAGCACCTCCCGGACGAGATGACTCAGCTGCCCATC	998
QY	2355	CAGCTTGTCCCTGTGCGGCCCCACTTTCCCGGGCTTAAGCAGCTGCTCCGCTGACCTTAA	2414
Db	999	CAGCTTGTCCCTGTGCGGCCCCACTTTCCCGGGCTTAAGCAGCTGCTCCGCTGACCTTAA	1058
QY	2415	AGACATCTCTAGCGAGGCCAGCACCATGCACTCCTTCAGCAACAGCAGCAGGAAGCAGT	2474
Db	1059	AGACATCTCTAGCGAGGCCAGCACCATGCACTCCTTCAGCAACAGCAGCAGGAAGCAGT	1118
QY	2475	ATCCGAAGCAGCAGCAGCGGGGAGCAGGAGGCGCTCGGGGGTCCCACTTCTCTCAA	2534
Db	1119	ATCCGAAGCAGCAGCAGCGGGGAGCAGGAGGCGCTCGGGGGTCCCACTTCTCTCAA	1178
QY	2535	GGACAATTACTTAGGGGGCACTTCGACCAATTTCTGACAAGCCAGAGGTGTGTGAAGC	2594
Db	1179	GGACAATTACTTAGGGGGCACTTCGACCAATTTCTGACAAGCCAGAGGTGTGTGAAGC	1238
QY	2595	AGTGTGCGTGTCCATGGCGCTGGTGTGAGCGCTTGGAGCATCTGAGTCCAGGGGAACA	2654
Db	1239	AGTGTGCGTGTCCATGGCGCTGGTGTGAGCGCTTGGAGCATCTGAGTCCAGGGGAACA	1298
QY	2655	GCTTCGGGGGATTTGATGTACGCCCACTTTTGGGAGTTCCACCGCTGTGCTGCCAC	2714
Db	1299	GCTTCGGGGGATTTGATGTACGCCCACTTTTGGGAGTTCCACCGCTGTGCTGCCAC	1358
QY	2715	TCCTTGTGCCCCATTTGGCCGAATGCAAGGTTCTCTGTCTAGCAGCAGCAGCAGGAAG	2774
Db	1359	TCCTTGTGCCCCATTTGGCCGAATGCAAGGTTCTCTGTCTAGCAGCAGCAGCAGGAAG	1418
QY	2775	CACCTGAAGTACTGCTGAGTATTTCCCTTTTCAGGGAGGTTTACACCAAGGCTAGAAG	2834
Db	1419	CACCTGAAGTACTGCTGAGTATTTCCCTTTTCAGGGAGGTTTACACCAAGGCTAGAAG	1478
QY	2835	CGAGAGCCTAGGCTGTCTGGCAGCGCTGCAGCAGGAGGCTCCGGGACACTTGAATGCC	2894
Db	1479	CGAGAGCCTAGGCTGTCTGGCAGCGCTGCAGCAGGAGGCTCCGGGACACTTGAATGCC	1538
QY	2895	GTCTACCTGTCTCTACAAGTCCGAGACCTGGAGGCGTGTGCTGTACAGAGTCG	2954
Db	1539	GTCTACCTGTCTCTACAAGTCCGAGACCTGGAGGCGTGTGCTGTACAGAGTCG	1598
QY	2955	CGACTACTACAACCTTTCCACTGGCTCTGGCGGAGCCGCCCTCCCTCCCGCTCCCA	3014

[illegible]

[illegible]

Db	2859	GTCCCGGATGTATACGCCAGTGTGTCCGAATGAGCGACCTCTCTCAAGAGTTTGGATGGCT	2918
QY	4266	CCAAATACCCCCCAGGAATTCCTGTGCATGAAGCACTGCTACTCTTTCACAGTATATTC	4325
Db	2919	CCAAATACCCCCCAGGAATTCCTGTGCATGAAGCACTGCTACTCTTTCACAGTATATTC	2978
QY	4326	AGTGGATGGCGTGAATAATCAAAATCTTTTGATGAACCTTCGAATGAATACATCAAGA	4385
Db	2979	AGTGGATGGCGTGAATAATCAAAATCTTTTGATGAACCTTCGAATGAATACATCAAGA	3038
QY	4386	ACTCGATCGTATCATGTGATGCACGAAAGAAAATCCACATCTGCTCAAGACGCTTCTA	4445
Db	3039	ACTCGATCGTATCATGTGATGCACGAAAGAAAATCCACATCTGCTCAAGACGCTTCTA	3098
QY	4446	CCAGCTCACCAAGCTCTGGACTCCGTCAGACCTATTGGAGAGAGCTGCATCAGTCCAC	4505
Db	3099	CCAGCTCACCAAGCTCTGGACTCCGTCAGACCTATTGGAGAGAGCTGCATCAGTCCAC	3158
QY	4506	TTTTGACCTGCTAATCAAGTCACACATGTCGAGCGTGGACTTCCGGAATGATGGCAGA	4565
Db	3159	TTTTGACCTGCTAATCAAGTCACACATGTCGAGCGTGGACTTCCGGAATGATGGCAGA	3218
QY	4566	GATCATCTCTGTCAAGTGCCCAAGATCCCTTCTCGGAAAGTCAAGCCCATCTATTCCA	4625
Db	3219	GATCATCTCTGTCAAGTGCCCAAGATCCCTTCTCGGAAAGTCAAGCCCATCTATTCCA	3278
QY	4626	CACCCAGTGAAGCATTTGGAAACCCATTTCCGCCACCCAGCTCATGCCCCCTTTCAGATG	4685
Db	3279	CACCCAGTGAAGCATTTGGAAACCCATTTCCGCCACCCAGCTCATGCCCCCTTTCAGATG	3338
QY	4686	TCTTCTGCCTGTTATAACTCTGCACCTACCTCTGCAGTGCCTTGGGGAATTCCTCTAT	4745
Db	3339	TCTTCTGCCTGTTATAACTCTGCACCTACCTCTGCAGTGCCTTGGGGAATTCCTCTAT	3398
QY	4746	TGATGTACAGTCTGTCATGAACATGTCCTGAATCTATTTCCTGGGCTTTTTCCTTC	4805
Db	3399	TGATGTACAGTCTGTCATG-----GAATCTATTTCCTGGGCTTTTTCCTTC	3447
QY	4806	TTTCTCTCCTCTTTTCTCTCTCTCCCTCCCTATCTACCCCTCCCATGGCCCTCAGAC	4865
Db	3448	TTTCTCTCCTCTTTTCTCTCTCTCCCTCCCTATCTACCCCTCCCATGGCCTTCAGAC	3507
QY	4866	TTTGCTTCCCATTTGGGCTCCCTATCTGTGTTTGAATGGTGTGTATGGCTTTAAATCTG	4925
Db	3508	TTTGCTTCCCATTTGGGCTCCCTATCTGTGTTTGAATGGTGTGTATGGCTTTAAATCTG	3567
QY	4926	TGATGATCCTCATATGGCCCATGTCAGTTGTGCTGTTTACAGCACTACTCTGTGCCA	4985
Db	3568	TGATGATCCTCATATGGCCCATGTCAGTTGTGCTGTTTACAGCACTACTCTGTGCCA	3627
QY	4986	GCCACACAAACGTTTACTTATCTTATGTCACGGGAAGTTTAGAGAGCTAAGATTCTGG	5045
Db	3628	GCCACACAAACGTTTACTTATCTTATGTCACGGGAAGTTTAGAGAGCTAAGATTCTGG	3687
QY	5046	GGAATCAAAACAAAAA	5062
Db	3688	GGAATCAAAACAAAAA	3704

[illegible]

QY	4866	TTTGGCTTCCCATTTGGCTCCTATCTCTGTTGTTTGAATGGTGTGTATGCCCTTTAAATCTG	4925
Db	3508	TTTGGCTTCCCATTTGGCTCCTATCTCTGTTGTTTGAATGGTGTGTATGCCCTTTAAATCTG	3567
QY	4926	TGATGATCCTCATATGCGCCAGTGTCAAGTTGCTTTTACACGACTACTCTGTGCCA	4985
Db	3568	TGATGATCCTCATATGCGCCAGTGTCAAGTTGCTTTTACACGACTACTCTGTGCCA	3627
QY	4986	GCACACAAACGTTTACTTATCTTATGCCACGGGAAGTTTACAGAGCTAAAGATATCTGG	5045
Db	3628	GCACACAAACGTTTACTTATCTTATGCCACGGGAAGTTTACAGAGCTAAAGATATCTGG	3687
QY	5046	GGAAATCAAAACAAAAA 5062	
Db	3688	GGAATCAAAACAAAAA 3704	
RESULT 5			
LOCUS	109479	3721 bp	DNA
DEFINITION	Sequence 1 from Patent WO 8909223.		
ACCESSION	109479		
VERSION	109479.1	GI:587808	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 3721)		
AUTHORS	Liao,S. and Chang,C.-S.		
TITLE	DNA BINDING PROTEINS INCLUDING ANDROGEN RECEPTOR		
JOURNAL	Patent: WO 8909223-A 1 05-OCT-1989;		
FEATURES	Location/Qualifiers		
source	1..3721		
BASE COUNT	843 a	1058 c	1004 g
ORIGIN	/organism="unknown"		
	815 t	1 others	
Query Match 70.3%; Score 3573.6; DB 6; Length 3721;			
Best Local Similarity 98.5%; Pred.No. 0;			
Matches 3687; Conservative 0; Mismatches 5; Indels 51; Gaps			
QY	1335	TGCACGGCGAGAGAACCTCTGTTTTCGCCCACTCTCTCCACCTCCTCCTGCCCTTCC	1394
Db	4	TTCCGGCGGAGAGAACCTCTGTTTTCGCCCACTCTCTCCACCTCCTCCTGCCCTTCC	63
QY	1395	CACCCCGAGTGGGAGCGGAGATCAAAAGATCAAAAGCGACTCAGGTCCTCAGTAGCCA	1454
Db	64	CACCCCGAGTGGGAG-CAGAGATCAAAAGATCAAAAGCGACTCAGGTCCTCAGTAGCCA	122
QY	1455	AAAAACAAACAAACAAAAACAAAAAGCGGAAATAAAAAGAAAAAGATAATAACTCAGTT	1514
Db	123	AAAAACAAACAAACAAAAACAAAAAGCGGAAATAAAAAGAAAAAGATAATAACTCAGTT	182
QY	1515	CTTATTGCACTTCTCAGTGACACTGAATTTGGAAGTGGAGATTTTGTGTTTTTTC	1574
Db	183	CTTATTGCACTTCTCAGTGACACTGAATTTGGAAGTGGAGATTTTGTGTTTTTTC	242
QY	1575	TTTTAAGATCTGGGCATCTTTTGAATCTACCCCTTCAAGTATTAAGAGACAGACTGTGAGC	1634
Db	243	TTTTAAGATCTGGGCATCTTTTGAATCTACCCCTTCAAGTATTAAGAGACAGACTGTGAGC	302
QY	1635	CTAGAGGGCAGATCTTGTCCACCGTGTGTCTTCTTCACAGAGACTTTGAGGCTGTCA	1694
Db	303	CTAGAGGGCAGATCTTGTCCACCGTGTGTCTTCTTCACAGAGACTTTGAGGCTGTCA	362
QY	1695	GAGCGCTTTTTCGGTGTGTCTCCCGCAAGTTTCTTCTGGAGTTCCCGCAGGTGGG	1754
Db	363	GAGCGCTTTTTCGGTGTGTCTCCCGCAAGTTTCTTCTGGAGTTCCCGCAGGTGGG	422
QY	1755	CAGTAGCTGCAGCGACTACCGCATCATCAGCGCTCTGCAACTCTTCTGACCAAGAAA	1814
Db	423	CAGTAGCTGCAGCGACTACCGCATCATCAGCGCTCTGCAACTCTTCTGACCAAGAAA	482

QY	1815	GGGAGCGGGGTAAAGGAAGTAGTGGAAAGATTACAGCAACCTCAAGATGGAAGTGA	1874
DB	483	GGGAGCGGGGGTAAAGGAAGTAGTGGAAAGATTACAGCAACCTCAAGATGGAAGTGA	542
QY	1875	GTTAGGGCTGGGAAGGGTCTACCCCTGGCCGCGCTCCAGACCTACCGAGGAGCTTTCCA	1934
DB	543	GTTAGGGCTGGGAAGGGTCTACCCCTGGCCGCGCTCCAGACCTACCGAGGAGCTTTCCA	602
QY	1935	GAATCTGTTCCAGAGCGTCGCGGAAGTGATCCAGAACCCGGGGCCCGAGCACCCAGAGGC	1994
DB	603	GAATCTGTTCCAGAGCGTCGCGGAAGTGATCCAGAACCCGGGGCCCGAGCACCCAGAGGC	662
QY	1995	CGCAGGCGCACACCTCCCGGCGCCAGTTTGCTGTGCTCTCAGCAGCAGCAGCAGCA	2054
DB	663	CGCAGGCGCACACCTCCCGGCGCCAGTTTGCTGTGCTCTCAGCAGCAGCAGCAGCA	701
QY	2055	GCACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC	2114
DB	702	---GCACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC	758
QY	2115	TAGCCCCCAGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG	2171
DB	759	TAGCCCCCAGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG	818
QY	2172	TAGAGCCCCACAGGCTACCTGGTCTCGATGAGGAACAGCAACCTTCACAGCGCAGTC	2231
DB	819	TAGAGCCCCACAGGCTACCTGGTCTCGATGAGGAACAGCAACCTTCACAGCGCAGTC	878
QY	2232	GGCCCTGGAAGTCCACCCCGAGAGAGTTGGTCCACAGAGCTTGGAGCGCGCTGCGCGC	2291
DB	879	GGCCCTGGAAGTCCACCCCGAGAGAGTTGGTCCACAGAGCTTGGAGCGCGCTGCGCGC	938
QY	2292	CAGCAAGGGGCTCGCGCAGCAGCTGCCAGCAGCTCCGAGCAGGATGACTCAGCTGCCCC	2351
DB	939	CAGCAAGGGGCTCGCGCAGCAGCTGCCAGCAGCTCCGAGCAGGATGACTCAGCTGCCCC	998
QY	2352	ATCCAGTTGTCCTCTGCTGGCGCCACATTTCCCGGCTTAAGCAGCTCTCGTGTGACCT	2411
DB	999	ATCCAGTTGTCCTCTGCTGGCGCCACATTTCCCGGCTTAAGCAGCTCTCGTGTGACCT	1058
QY	2412	TAAAGACATCTGAGCGAGGCCAGCACCATGCACTCTTAGCAA---CAGCAGCAGGA	2468
DB	1059	TAAAGACATCTGAGCGAGGCCAGCACCATGCACTCTTAGCAA---CAGCAGCAGGA	1118
QY	2469	AGCAGTATCCGAAAGCAGCAGCAGCGGGAGAGCGAGGAGGCGCTCGGGGGCTCCCACTTC	2528
DB	1119	AGCAGTATCCGAAAGCAGCAGCAGCGGGAGAGCGAGGAGGCGCTCGGGGGCTCCCACTTC	1178
QY	2529	CTCCAGGACAAATTACTTAGGGGCACTTCGACCATTCTGACACGCCAAGGACTGTGG	2588
DB	1179	CTCCAGGACAAATTACTTAGGGGCACTTCGACCATTCTGACACGCCAAGGACTGTGG	1238
QY	2589	TAGGCGAGTCCGGTGCCATGGCTGGGTGTGAGGGGTTGGAGCATCTGAGTCCAGG	2648
DB	1239	TAGGCGAGTCCGGTGCCATGGCTGGGTGTGAGGGGTTGGAGCATCTGAGTCCAGG	1298
QY	2649	GGAACAGCTTCGGGGGATGTCATGTACGCCCCACTTTTGGGAGTTCCACCGCTGTGG	2708
DB	1299	GGAACAGCTTCGGGGGATGTCATGTACGCCCCACTTTTGGGAGTTCCACCGCTGTGG	1358
QY	2709	TCCCACTCTTGTGCCCATTTGGCCGAATGCAAGGTCTCTCTCTAGACGACAGCCGAGG	2768
DB	1359	TCCCACTCTTGTGCCCATTTGGCCGAATGCAAGGTCTCTCTCTAGACGACAGCCGAGG	1418
QY	2769	CAAGAGCACTGAAGATACCTGCTAGTATTTCCCTTTCAAGGGAGGTTTACACCAAGGGCT	2828
DB	1419	CAAGAGCACTGAAGATACCTGCTAGTATTTCCCTTTCAAGGGAGGTTTACACCAAGGGCT	1478
QY	2829	AGAAGCGCAGACCTTAGCTCTCTTGCGAGCGCTGACAGCGGAGGAGCTCCGGGACACTTGA	2888
DB	1479	AGAAGCGCAGACCTTAGCTCTCTTGCGAGCGCTGACAGCGGAGGAGCTCCGGGACACTTGA	1538
QY	2889	ACTGCCGCTACCCCTGCTCTCTACAGTCCGGGACACTGGACGAGCAGCTGCGTACCA	2948

Db 1539 ACTGCCGTACCTGCTCTCTACAGTCCGAGCACTGGAGGAGCTGCGTACCA 1598
 QY 2949 GAGTCGGCACTACTACAATTCACACTGCTCTGCGCGAGCGCGCGCCCTCGCGCGC 3008
 Db 1599 GAGTCGGCACTACTACAATTCACACTGCTCTGCGCGAGCGCGCGCCCTCGCGCGC 1658
 QY 3009 TCCCACTCCCACTGCGCTCAAGCTGGAGAACCGCTGAGCTACGCGAGCGCTGGGC 3068
 Db 1659 TCCCACTCCCACTGCGCTCAAGCTGGAGAACCGCTGAGCTACGCGAGCGCTGGGC 1718
 QY 3069 GCGTCGGCGCGCGAGTCCGCTATGGGACCTGCGAGCGCTGATGCGCGGCTGCGC 3128
 Db 1719 GCGTCGGCGCGCGAGTCCGCTATGGGACCTGCGAGCGCTGATGCGCGGCTGCGC 1778
 QY 3129 GGGACCGGCTTGGGTCAACCTCAGCGCGCGCTTCCCTATCCCTGGCACACTCTCTTAC 3188
 Db 1779 GGGACCGGCTTGGGTCAACCTCAGCGCGCGCTTCCCTATCCCTGGCACACTCTCTTAC 1838
 QY 3189 AGCGGAAGAGCCAGTTGTATGACCGCTGTGGTGGTGGGCTGGT-----GG 3239
 Db 1839 AGCGGAAGAGCCAGTTGTATGACCGCTGTGGTGGTGGGCTGGTGGCGCGG 1898
 QY 3240 GCG 3299
 Db 1899 GCG 1958
 QY 3300 AGCTGTAGCCCTACCGCTACACTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3359
 Db 1959 AGCTGTAGCCCTACCGCTACACTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2018
 QY 3360 CTTACCGCACTGATGT 3419
 Db 2019 CTTACCGCACTGATGT 2078
 QY 3420 TCCCACTGTGTCAAGAGCGAATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3479
 Db 2079 TCCCACTGTGTCAAGAGCGAATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2138
 QY 3480 GGACATGCGTTGGAGACTGCCAGGACCATGTTTGGCCATGTACTATGACTTATTTCCACC 3539
 Db 2139 GGACATGCGTTGGAGACTGCCAGGACCATGTTTGGCCATGTACTATGACTTATTTCCACC 2198
 QY 3540 CCAGAGACTGCTGATCTGTGGAGTGAAGCTTCTGGGTGTCACACTGAGGCTCTCAC 3599
 Db 2199 CCAGAGACTGCTGATCTGTGGAGTGAAGCTTCTGGGTGTCACACTGAGGCTCTCAC 2258
 QY 3600 ATGTGAAGCTGCAAGGTCTTCTCAAAAGAGCGCTCAAGGGGAAACAGAGTACCTGTG 3659
 Db 2259 ATGTGAAGCTGCAAGGTCTTCTCAAAAGAGCGCTCAAGGGGAAACAGAGTACCTGTG 2318
 QY 3660 GCGCAGCAAGATGATGCACTATTTGATAATTCGAGGAAATTTGTCACCTGTGCG 3719
 Db 2319 GCGCAGCAAGATGATGCACTATTTGATAATTCGAGGAAATTTGTCACCTGTGCG 2378
 QY 3720 TCTCGGAAATGTTATGAAGCAGGATGACTCTGGAGCGCGGAGCTGAAGAACTTTGG 3779
 Db 2379 TCTCGGAAATGTTATGAAGCAGGATGACTCTGGAGCGCGGAGCTGAAGAACTTTGG 2438
 QY 3780 TAATCTGAATACAGGAGGAGGAGGCTTCCAGCACCACCGCGCGCGCGCGCGCGCGCG 3839
 Db 2439 TAATCTGAATACAGGAGGAGGAGGCTTCCAGCACCACCGCGCGCGCGCGCGCGCGCG 2498
 QY 3840 AACCCAGAGCTGACAGTGTACACATTTGAAGCTATGAATGTACGCCATCTTTCTGAA 3899
 Db 2499 AACCCAGAGCTGACAGTGTACACATTTGAAGCTATGAATGTACGCCATCTTTCTGAA 2558
 QY 3900 TGTCTGGAAGCCATTGAGCAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3959
 Db 2559 TGTCTGGAAGCCATTGAGCAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2618
 QY 3960 CTCCTTTGAGCCTTCTCTAGCTCAATGAACCTGGAGAGGAGCTTGTACAGT 4019

Db 2619 CTCTTTGAGCGCTTGTCTTAGCTCAATGAACCTGGAGAGACAGCTTGTACAGT 2678
 QY 4020 GGTCAAGTGGCGCAAGCGCTTGCCTGGCTTCGGAACCTTACAGTGGAGACGACGATGGC 4079
 Db 2679 GGTCAAGTGGCGCAAGCGCTTGCCTGGCTTCGGAACCTTACAGTGGAGACGACGATGGC 2738
 QY 4080 TGTCAATCAGTACTCTCTGATGGGCTCATGCTGTTTGGCCATGGCTGGCGATCCTTCAC 4139
 Db 2739 TGTCAATCAGTACTCTCTGATGGGCTCATGCTGTTTGGCCATGGCTGGCGATCCTTCAC 2798
 QY 4140 CAATGTCAACTCCAGGATGCTCTACTTCCGCCCTGATCTGTTTAAATGAGTACCCAT 4199
 Db 2799 CAATGTCAACTCCAGGATGCTCTACTTCCGCCCTGATCTGTTTAAATGAGTACCCAT 2858
 QY 4200 GCACAAGTCCCGGATGTACAGCCAGTGTCCGAATGAGGCACCTCTCTCAAGAGTTGG 4259
 Db 2859 GCACAAGTCCCGGATGTACAGCCAGTGTCCGAATGAGGCACCTCTCTCAAGAGTTGG 2918
 QY 4260 ATGGCTCAAAATCACCCCGGAGGATTCCTGTGCATGAAAGCACTGCTACTTTCAGCAT 4319
 Db 2919 ATGGCTCAAAATCACCCCGGAGGATTCCTGTGCATGAAAGCACTGCTACTTTCAGCAT 2978
 QY 4320 TATTCAGTGTGATGGCTGAAATCAAAATTTCTTGTGAACTTTCGAATGAACATACAT 4379
 Db 2979 TATTCAGTGTGATGGCTGAAATCAAAATTTCTTGTGAACTTTCGAATGAACATACAT 3038
 QY 4380 CAAGGAACCTGATCGTATCATTTGATGCAAAAGAAATCCACATCTCTGTCAGAG 4439
 Db 3039 CAAGGAACCTGATCGTATCATTTGATGCAAAAGAAATCCACATCTCTGTCAGAG 3098
 QY 4440 CTTCTACCGCTCACCAAGCTCTGAGTCCGCTGAGCCTATTGGAGAGAGCTGCATCA 3158
 Db 3099 CTTCTACCGCTCACCAAGCTCTGAGTCCGCTGAGCCTATTGGAGAGAGCTGCATCA 4499
 QY 4500 GTTCACTTTTGAAGCTGATCAAGTCAACATGTCAGCTGAGCTGAGCTTTCGGAATGAT 4559
 Db 3159 GTTCACTTTTGAAGCTGATCAAGTCAACATGTCAGCTGAGCTGAGCTTTCGGAATGAT 3218
 QY 4560 GGCAGAGATCATCTCTGTGCAAGTGCCTGAGTCCCTTCTGGAAAGTCAAGCCATCTCA 4619
 Db 3219 GGCAGAGATCATCTCTGTGCAAGTGCCTGAGTCCCTTCTGGAAAGTCAAGCCATCTCA 3278
 QY 4620 TTTCCACACCGCTGAGTCAAGTGCCTGAGTCCCTTCTGGAAAGTCAAGCCATCTCA 4679
 Db 3279 TTTCCACACCGCTGAGTCAAGTGCCTGAGTCCCTTCTGGAAAGTCAAGCCATCTCA 3338
 QY 4680 CAGATGCTTCTGCTGCTGATTAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4739
 Db 3339 CAGATGCTTCTGCTGCTGATTAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3398
 QY 4740 CTTCTATGATGTACAGTCTGTCATGACATGCTTCTGATTTGCTTATTTGCTTATTTT 4799
 Db 3399 CTTCTATGATGTACAGTCTGTCATG-----GAATTTCTATTTGCTTATTTT 3447
 QY 4800 TTTCTTCTTCT 4859
 Db 3448 TTTCTTCTTCT 3507
 QY 4860 TCAGACTTTGCTTCT 4919
 Db 3508 TCAGACTTTGCTTCT 3567
 QY 4920 ATCTGTGATGATCTCTATATGCGGAGTCTCAAGTGTGCTGCTGCTGCTGCTGCTGCT 4979
 Db 3568 ATCTGTGATGATCTCTATATGCGGAGTCTCAAGTGTGCTGCTGCTGCTGCTGCTGCT 3627
 QY 4980 GTCCAGCCACACAAAGTTTACTTATGTCACGGAAGTTTAGAGAGCTAAGATT 5039
 Db 3628 GTCCAGCCACACAAAGTTTACTTATGTCACGGAAGTTTAGAGAGCTAAGATT 5099
 QY 5040 ATCTGGGAAATCAAAACAAAAA 5062
 Db 3688 ATCTGGGAAATCAAAACAAAAA 3710

[illegible]

```
QY 3302 CTGTAGCCCTTACGGCTACACTGGCCCTCAGGGCTGGCGGCAGGAAGCGACT 3361
Db 1789 TTGTAGCCCTTACGGCTACACTGGCCCTCAGGGCTGGCGGCAGGAAGCGACT 1848
QY 3362 TCACCGACCTGATGTGTGTACCTGGCGGATGTGTGAGCAGAGTGCCTTATCCAGTTC 3421
Db 1849 TCACCGACCTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1908
QY 3422 CCATCTGTGTCAAAGGAAATGGCCCTGGGATGATGATGATGATGATGATGATGATG 3481
Db 1909 CCATCTGTGTCAAAGGAAATGGCCCTGGGATGATGATGATGATGATGATGATGATG 3481
QY 3482 ACATGCGCTTGGAGACTGCCAGGACCATGTTTGGCCATGATGATGATGATGATGATG 3541
Db 1969 ACATGCGCTTGGAGACTGCCAGGACCATGTTTGGCCATGATGATGATGATGATGATG 2028
QY 3542 AGAAGACCTGCTGATCTGTGAGATGAAGCTTCTGGGTGTGACATGATGAGCTCTCAT 3601
Db 2029 AGAAGACCTGCTGATCTGTGAGATGAAGCTTCTGGGTGTGACATGATGAGCTCTCAT 2088
QY 3602 GTGGAAGCTGCAAGTCTTCTTCAAAGAGCCGCTGAAGGAAACAGAAAGTACCTGTGCG 3661
Db 2089 GTGGAAGCTGCAAGTCTTCTTCAAAGAGCCGCTGAAGGAAACAGAAAGTACCTGTGCG 2148
QY 3662 CCAGCAAAATGATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3721
Db 2149 CCAGCAAAATGATGTCATGATGATGATGATGATGATGATGATGATGATGATGATG 3721
QY 3722 TTGCGAAATGTTATGAAGAGGATGACTCTGGAGCCCGGAAGCTGAAGAACTTGGTA 2208
Db 2209 TTGCGAAATGTTATGAAGAGGATGACTCTGGAGCCCGGAAGCTGAAGAACTTGGTA 2268
QY 3782 ATCTGAACTACAGAGGAGGAGGCTTCCAGCACACACAGCCGCCACTGAGGAGACAA 3841
Db 2269 ATCTGAACTACAGAGGAGGAGGCTTCCAGCACACACAGCCGCCACTGAGGAGACAA 2328
QY 3842 CCCAAGAGCTGACAGTGTACACATGTAAGCTATGAATGTACAGCCCATCTTCTGAATG 3901
Db 2329 CCCAAGAGCTGACAGTGTACACATGTAAGCTATGAATGTACAGCCCATCTTCTGAATG 3901
QY 3902 TCTTGAAGCCATGAGCAGGCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3961
Db 2389 TCTTGAAGCCATGAGCAGGCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2448
QY 3962 CTTTGTAGCCTTCTCTAGCCTCAATGAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAG 4021
Db 2449 CTTTGTAGCCTTCTCTAGCCTCAATGAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAG 2508
QY 4022 TCAAGTGGCCCAAGGCTTCCCTGGCTTCCGCAACTTACAGTGGAGACAGAGAGAGAGAG 4081
Db 2509 TCAAGTGGCCCAAGGCTTCCCTGGCTTCCGCAACTTACAGTGGAGACAGAGAGAGAGAG 2568
QY 4082 TCATTTCAGTACTCTGTGAGGAGCTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4141
Db 2569 TCATTTCAGTACTCTGTGAGGAGCTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2628
QY 4142 ATGTCAACTCCAGAGTCTTACTTCCGCCCTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4201
Db 2629 ATGTCAACTCCAGAGTCTTACTTCCGCCCTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2688
QY 4202 ACAAGTCCCGATGTACAGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4261
Db 2689 ACAAGTCCCGATGTACAGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2748
QY 4262 GGTCCCAATACCCCGAGGATTCCTGTGATGAAAGCACTGTCTTCTTTCAGCATTA 4321
Db 2749 GGTCCCAATACCCCGAGGATTCCTGTGATGAAAGCACTGTCTTCTTTCAGCATTA 2808
QY 4322 TTCCAGTGGTGGTGAAGAAATCAAAATCTTGTGATGAACTTGTGATGAACTTGTGATGAA 4381
Db 2809 TTCCAGTGGTGGTGAAGAAATCAAAATCTTGTGATGAACTTGTGATGAACTTGTGATGAA 2868
QY 4382 AGGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4441
```

```
Db 2869 AGGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2928
QY 4442 TCTACAGCTACCAAGCTCCTGGACTCCGTCAGCCTATTGCGAGAGAGCTGCAATCAGT 4501
Db 2929 TCTACAGCTACCAAGCTCCTGGACTCCGTCAGCCTATTGCGAGAGAGCTGCAATCAGT 2988
QY 4502 TCACTTTTGTACCTGCTTAATCAAGTCAACATGTTGAGCGTGGAGCTTCCGGAATGATGG 4561
Db 2989 TCACTTTTGTACCTGCTTAATCAAGTCAACATGTTGAGCGTGGAGCTTCCGGAATGATGG 3048
QY 4562 CAGAGATCATCTGTGCAAGTGCCTTCTGGGAAAGTCAAGCCCACTATT 4621
Db 3049 CAGAGATCATCTGTGCAAGTGCCTTCTGGGAAAGTCAAGCCCACTATT 3108
QY 4622 TCCACACCCAGTGAAGCATTTGAAACCCCTATTTCCTCCACCCAGCTCATGCCCCTTCA 4681
Db 3109 TCCACACCCAGTGAAGCATTTGAAACCCCTATTTCCTCCACCCAGCTCATGCCCCTTCA 3168
QY 4682 GATGCTTCTGCTGTTAATCACTCTGCACTACTCTCTGCACTGCTTGGGAAATTTTCT 4741
Db 3169 GATGCTTCTGCTGTTAATCACTCTGCACTACTCTCTGCACTGCTTGGGAAATTTTCT 3228
QY 4742 CTATTGATCTACAGTCTGTCAATGAACATGTTCTGCTGCTGCTGCTGCTGCTGCTGCT 4801
Db 3229 CTATTGATCTACAGTCTGTCAATGAACATGTTCTGCTGCTGCTGCTGCTGCTGCTGCT 3288
QY 4802 TCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4861
Db 3289 TCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3348
QY 4862 AGACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4921
Db 3349 AGACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3408
QY 4922 TCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4981
Db 3409 TCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3468
QY 4982 GCCAGCCACACAAAGCTTACTTATCTATGTCACGCGGAGTGTAGAGAGCTAAGATTAT 5041
Db 3469 GCCAGCCACACAAAGCTTACTTATCTATGTCACGCGGAGTGTAGAGAGCTAAGATTAT 3528
QY 5042 CTGGGAAATCAAAACAAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 5082
Db 3529 CTGGGAAATCAAAACAAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 3569
```

RESULT 7

HUMANDEC

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

FEATURES

source

3569 bp mRNA linear PRI 31-OCT-1994
Human androgen receptor (AR) mRNA, complete cds.
M20132 J03180
M20132.1 GI:178627
androgen receptor.
Human, epididymal cDNA to mRNA, clones ARHEL[1-3] and ARHEL1.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
Lubahn, D.B., Joseph, D.R., Sar, M., Tan, J., Higgs, H.N., Larson, R.E.,
French, F.S. and Wilson, E.M.
The human androgen receptor: complementary deoxyribonucleic acid
cloning, sequence analysis and gene expression in prostate
Mol. Endocrinol. 2 (12), 1265-1275 (1988)
89112208
3216866
Draft entry and computer readable sequence [1] kindly submitted by
E.M.Wilson, 18-AUG-1988.
Location/Qualifiers
1..3569
/organism="Homo sapiens"
/db_xref="taxon:9606"

gene	CD5	BASE COUNT	796 a	1009 c	974 g	790 t
/map="Xq11.2-q12"						
1. .3569						
/gene="AR"						
363. .3122						
/gene="AR"						
/note="androgen receptor"						
/codon_start=1						
/protein_id="AA511729.1"						
/db_xref="GI:178628"						
/db_xref="GDB:100-120-556"						
/translation="MEVQLGLGRVYPPRPSKYRGAFQNLFSVREVIONPGRHPDEA						
ASAPPGASLLLLQOOQQ						
TGYVLDEEQPQSPQSALECHPERGCPVPGFAGAAKGLKQOLPAPDEDDSAAPS						
TSLSLGTFPGLSKDLILSEASTMOLLOQOOQAAVESSSSGRAREAGAPTS						
SKDPTLGGTSTISDNKELCKAVSNGLVEALESPGQELRGQDCTTAPLLGVDA						
VRNTPCAPLACKKSLDSDSAGSTEDTAETSPFKGTGKLESGSGGSAAGSS						
GTGLPSTLSLYKSGDALHCAAGAPGSGPSAAASSSWHTLFTAEQGLVPGCG						
GSGSWAAAAOORRYKGLDILHCAAGAPGSGPSAAASSSWHTLFTAEQGLVPGCG						
GG						
SWKPTSPICVYKSEMGWMDVSQCPYGMRLTARDHVLIDYVFPQKTLCTGDEA						
SCSHGATLVCKSGYFKRAEGKQVLCASRNDCTDKPRKNCPURLKCYEAGM						
TLGARLKLKGLNKLDEEGEASLSTETETQKLTYSHLGVECOPIFLNLEALEP						
GYVCAGHDNNQDLPAAALLNSLSEITGELQVLVHVWAKALPGPNLHDDQMAVIOY						
WMLGMYFAMGRNFTNVMSRLITAPDLVFNFEYRMKSRMISQCVNMRHLKTSOEFGLQ						
ITPQEFLLMKALLSITPVDGLKPKCFDELNRNYSIKELDIRACKRHLNPTSCSRRF						
YQLTKLLDSQVPIARELHQFTFDLLIKSHMYSVDFPEMMAEIIISVQVFKILSKVAPI						
YFHTQ"						
QY	1502	TAATAACTCAGTCTCTATTATTTGCACCTACTTCACTGACGACACTGAATTTGGAAGGTGAGGA	1561			
DB	1	TATTAATCTCAGTCTCTATTATTTGCACCTACTTCACTGACGACACTGAATTTGGAAGGTGAGGA	60			
QY	1562	TTTTCTGTTTTTCTTTTAAGATCTGGGCATCTTTTGAATCTACCCCTCAAGTATTAAAG	1621			
DB	61	TTTTCTGTTTTTCTTTTAAGATCTGGGCATCTTTTGAATCTACCCCTCAAGTATTAAAG	120			
QY	1622	ACAGACTGTGAGCCTTAGCAGGCGAGATCTTGTCCACCGTGTGCTTCTGTGCACGAGAC	1681			
DB	121	ACAGACTGTGAGCCTTAGCAGGCGAGATCTTGTCCACCGTGTGCTTCTGTGCACGAGAC	180			
QY	1682	TTTGAGGCTGTGAGAGCGCTTTTTCGTTGGTGTCTCCCGAAGTTTCCTTCTCTGAGCT	1741			
DB	181	TTTGAGGCTGTGAGAGCGCTTTTTCGTTGGTGTCTCCCGAAGTTTCCTTCTCTGAGCT	240			
QY	1742	TCCCGCAGGTGGCAGCTAGCTGTCAGGCACCTACCGCATCATCAGAGCTTTTGAATCTTT	1801			
DB	241	TCCCGCAGGTGGCAGCTAGCTGTCAGGCACCTACCGCATCATCAGAGCTTTTGAATCTTT	300			
QY	1802	CTGAGCAAGAGAGGGAGCGGGGTAGGGAAGTAGTGGAAGATTTCAGCAAGCTCAA	1861			
DB	301	CTGAGCAAGAGAGGGAGCGGGGTAGGGAAGTAGTGGAAGATTTCAGCAAGCTCAA	360			
QY	1862	GGATGGAAGTGCAGTTAGGGCTGGGAAGGTCTACCCCTCGCGCGCTTCAAGACCTACC	1921			
DB	361	GGATGGAAGTGCAGTTAGGGCTGGGAAGGTCTACCCCTCGCGCGCTTCAAGACCTACC	420			
QY	1922	GAGAGCTTTTCCAGATCTGTTCCAGAGCGTGCAGGAAGTAGATCCAGAACCCCGGCCCA	1981			
DB	421	GAGAGCTTTTCCAGATCTGTTCCAGAGCGTGCAGGAAGTAGATCCAGAACCCCGGCCCA	480			
QY	1982	GGCACCCAGAGCGCGGAGCGCAGCACCTCCCGGGGCCAGTTGTGTGTGTGCAGCAGC	2041			
DB	481	GGCACCCAGAGCGCGGAGCGCAGCACCTCCCGGGGCCAGTTGTGTGTGTGCAGCAGC	540			
QY	2042	AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	2101			
DB	541	AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	598			

[illegible]

Db	2749	GGCTCAAATACCCCCAGGAATTCCTGTGCATGAAGCACTGCTACTCTCTCAACATTA	2808
QY	4322	TTCCAGTGGATGGCTGAAAAATCAAAAATCTTTTGATGAACCTCGAATGAACATACATCA	4381
Db	2809	TTCCAGTGCATGGCTGAAAAATCAAAAATCTTTTGATGAACCTCGAATGAACATACATCA	2868
QY	4382	AGGAACCTGCATCGTATCATTCGATGCAAAAGAAAAAATCCACACATCTCTGCTCAAGACGCT	4441
Db	2869	AGGAACCTGCATCGTATCATTCGATGCAAAAGAAAAAATCCACACATCTCTGCTCAAGACGCT	2928
QY	4442	TCTACACAGCTACCAAGCTCCTGGACTCCGTGCGAGCCTATTGCGAGAGAGCTGCATCAGT	4501
Db	2929	TCTACACAGCTACCAAGCTCCTGGACTCCGTGCGAGCCTATTGCGAGAGAGCTGCATCAGT	2988
QY	4502	TCACHTTTGACCTGCTAATCAAGTCACACATGGTGAGCGTGGACTTTCCGGAAATCATGG	4561
Db	2989	TCACHTTTGACCTGCTAATCAAGTCACACATGGTGAGCGTGGACTTTCCGGAAATCATGG	3048
QY	4562	CAGAGATCATCTCTGTGCAAGTGCCCAAGATCCCTTCTGGGAAGTCAAGCCCATCTATT	4621
Db	3049	CAGAGATCATCTCTGTGCAAGTGCCCAAGATCCCTTCTGGGAAGTCAAGCCCATCTATT	3108
QY	4622	TCCACACCCAGTGAAGCATTTGGAACCCCTATTTCGCCACCCAGCTCATGCCCTTTCA	4681
Db	3109	TCCACACCCAGTGAAGCATTTGGAACCCCTATTTCGCCACCCAGCTCATGCCCTTTCA	3168
QY	4682	GATGCTCTTCGCCGTTTATAAATCTGCACCTACTCTCTCGAGTGCCCTTGGGAAATTTCT	4741
Db	3169	GATGCTCTTCGCCGTTTATAAATCTGCACCTACTCTCTCGAGTGCCCTTGGGAAATTTCT	3228
QY	4742	CTATTGATGACAGTCTGTGCATGAACATGTTCTCTGAATTCATTTCGTGCGCTTTTCTTT	4801
Db	3229	CTATTGATGACAGTCTGTGCATGAACATGTTCTCTGAATTCATTTCGTGCGCTTTTCTTT	3288
QY	4802	TCCTCTTCTCTCTCTTTTCTTCTCCCTCCTCATCAACCTCCCATGCCACCTTC	4861
Db	3289	TCCTCTTCTCTCTCTTTTCTTCTCCCTCCTCATCAACCTCCCATGCCACCTTC	3348
QY	4862	AGACTTTGCTTCCATTGGGCTCCTCATCTGTGTTTGAATGGTCTCTATGCGCTTTAAA	4921
Db	3349	AGACTTTGCTTCCATTGGGCTCCTCATCTGTGTTTGAATGGTCTCTATGCGCTTTAAA	3408
QY	4922	TCGTGCATCATCTCATATGGCCCGAGTGCAAGTTGTGCTTTTACAGCACTACTCTGT	4981
Db	3409	TCGTGCATCATCTCATATGGCCCGAGTGCAAGTTGTGCTTTTACAGCACTACTCTGT	3468
QY	4982	GCCAGCCACACAAGGTTTACTTATCTTATGCCAGGGGAAGTTTAGAGACCTATGATTAT	5041
Db	3469	GCCAGCCACACAAGGTTTACTTATCTTATGCCAGGGGAAGTTTAGAGACCTATGATTAT	3528
QY	5042	CTGGGGAATCAAAACAAAAACAAACAAACAAAAA	5082
Db	3529	CTGGGGAATCAAAACAAAAACAAACAAACAAAAA	3569

RESULT	8
HSDJ	0804
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	

HSDJ80804 139033 bp DNA linear PRI 09-MAR-2002
Human DNA sequence from clone RP4-80804 on chromosome Xq11.1-12,
complete sequence.
AL049564
AL049564.11 GI:19338177
HTG.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 139033)
Chapman, J.
Direct Submission
Submitted (08-MAR-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

COMMENT

humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Mar 10, 2002 this sequence version replaced gi:4902757.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome X, constructed by the Sanger Centre Chromosome X Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/ChrX
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. RP4-80804 is from the
library RP4-4 constructed by the group of Pieter de Jong. For
further details see
http://www.chori.org/bacpac/home.htm
Vektor: pCYPAC2
This sequence is the entire insert of clone RP4-80804 The true
right end of clone RP6-22p16 is at 47989 in this sequence.

FEATURES

source

Location/Qualifiers
1. 139033
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="q11.1-12"
/clone="RP4-80804"
/clone_lib="RP4-4"

BASE COUNT 46773 a 25978 c 26907 g 39375 t

Query Match 62.7%; Score 3185.4; DB 9; Length 139033;
Best Local Similarity 97.7%; Pred. No. 0;
Matches 3428; Conservative 0; Mismatches 40; Indels 40; Gaps 19;
QY 1 GAGCTCTGACAAAATTGAGCGCCTATCTGTACATGGCAAGTGTCTTACTGTGTGTG 60
Db 102221 GAGCTCTGACAAAATTGAGCGCCTATGTGTACATGGCAAGTGTCTTACTGTGTGTG 102280
QY 61 TTTACCTGCTTGTCTGGGTGATTTTGCCTTTGAGAGTCTGGATGAGAAATGCGTAA 120
Db 102281 TTTACCTGCTTGTCTGGGTGATTTTGCCTTTGAGAGTCTGGATGAGAAATGCGTAA 102340
QY 121 AGGCAATTCAGACAGGAAGAGCAGAGAGAGGGTGAAGATGACCTCTGATCTTGG 180
Db 102341 AGGCAATTCAGACAGGAAGAGCAGAGAGAGGGTGAAGATGACCTCTGATCTTGG 102400
QY 181 GGCTGAGGGTCTCAGAGCAATGGCACAATGCCAGAGGCCGATCTATCCCTATGCG 240
Db 102401 GGCTGAGGGTCTCAGAGCAATGGCACAATGCCAGAGGCCGATCTATCCCTATGCG 102460
QY 241 GAACCTAAGGTTTACGATCAGCTATCTGCTGGCTTGGTCTACTGGCTTGGCTCTCAGT 300
Db 102461 GAA-TCTAAGGTTTACG--CAAGTATCTGCTGGCTTGGTCTA-TGGCTTGGCTCTCAGT 102514
QY 301 TTGTAGGAGACTCTCCCACTCTCCCATCTGCGGCTCTTATCAGTCTTCTGAAAGAACCCN 360
Db 102515 TTGTAGGAGACTCTCCCACTCTCCCATCTGCGGCTCTTATCAGTCTTCTGAAAGAACCC 102574
QY 361 TGGNAGCCAGGACNAGGATTCNTATCGTCTCTTTCTCTCTCTCTCTCTCTCTCTCTCT 416
Db 102575 TGGC-AGCCAGGAGC-AGGATTCCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 102632

QY 417 NGTTGNTTTTAGATTGCTTNGNAACCAAAATT---TGTATGCTGGCTCCAGGAATC 473
Db 102633 GTTGTGTTTTTAGATTGGCTTTTGGACCAAAATTGTTGAGTGTCTGGCTCCAGGAATC 102692
QY 474 TGGAG-CCTGCGCGCTAAACCTTGTGTTAGAAAGCAGGAGCTATTTCAGAAACA-GGCT 531
Db 102693 TGGAGCCTCGCGCTAAACCTTGTGTTAGAAAGCAGGAGCTATTTCAGAAACAGGGGT 102752
QY 532 CTTCCAGGCTTAGAGTAGCTTCTCTGCTTCTCTGCTTCTCTGCTTCTCTGCTTCTCT 590
Db 102753 CTTCCAGGCTTAGAGTAGCTTCTCTGCTTCTCTGCTTCTCTGCTTCTCTGCTTCTCT 102812
QY 591 CCAAGACNACTAGGAGCGCTTGTAGCGCGGTGAGGGAGGAGAAAGAAAGGGAG 650
Db 102813 CCAAGACNACTAGGAGCGCTTGTAGCGCGGTGAGGGAGGAGAAAGAAAGGGAG 102872
QY 651 GGGAGGAAAGAGGTGGGAGGCAAGGAGCGCGGCTGGGGGGGGGGGGGGGGGGGGGG 710
Db 102873 GGGAGGAAAGAGGTGGGAGGCAAGGAGCGCGGCTGGGGGGGGGGGGGGGGGGGGGG 102932
QY 711 GCANNAACTGTTGCAATTTGCTTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 770
Db 102933 GCA--AACTGTTGCAATTTGCTTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 102990
QY 771 AGCTTCTGGAGAGCGGAGCGGTCCGAGCAAGCCCGAGAGGAGGAGGAGGAGGAGGAG 830
Db 102991 AGCTTCTGGAGAGCGGG-ACGCTCCGAGCAAGCCCGAGAGGAGGAGGAGGAGGAG 103049
QY 831 GGAAGAGGCGCCNAGCTAGCCGCTCCAGTGTGTACAGNAGCCGAA--GGAGCAGCACG 889
Db 103050 GGAAGAGG-CCGAGTAGCCGCTCCAGTGTGTACAGNAGCCGAAAGGAGGAGGAGGAG 103108
QY 890 CCAGCCCGAGCGGCTCCAGGAGCAGACGAGCCTCTTTCGA-----NGCTTTCGAAGC 943
Db 103109 CCAGCCCGAGCGGCTCCAGGAGCAGACGAGCCTCTTTCGAAGCAGCGCGGCTTCGAAGC 103168
QY 944 CGCGCCCGGAGCTGCGCTTCTCTTCTCGTGAAGTGTGTTTAAAGCTGCTAAGACTCGG 1003
Db 103169 CGCGCCCGGAGCTGCGCTTCTCTTCTCGTGAAGTGTGTTTAAAGCTGCTAAGACTCGG 103228
QY 1004 AGGAAGCAAGAAAGTGCCTTGTAGGACTGAGCGCTTGTCTCTCTCTCTCTCTCTCTCT 1063
Db 103229 AGGAAGCAAGAAAGTGCCTTGTAGGACTGAGCGCTTGTCTCTCTCTCTCTCTCTCTCT 103288
QY 1064 CGCGCTCCCGCCAGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1123
Db 103289 CGCGCTCCCGCCAGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 103348
QY 1124 TGGGTACTCTCAGCAACCCCTCAGCAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1183
Db 103349 TGGGTACTCTCAGCAACCCCTCAGCAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 103408
QY 1184 TCGGCCAGCGCTGCAAGNCCGAGTTTTCAGAGAGGTAACTCCCTTGGCTTCGAGCGGG 1243
Db 103409 TCGGCCAGCGCTGCAAGNCCGAGTTTTCAGAGAGGTAACTCCCTTGGCTTCGAGCGGG 103468
QY 1244 CGAGNCTAGCTGCACATTTGCAAGAGGCTCTTAGGAG-CAGGCGACTTGGGAGCGGCTT 1302
Db 103469 CGAG-CTAGCTGCACATTTGCAAGAGGCTCTTAGGAGCGGCGACTTGGGAGCGGCTT 103527
QY 1303 CAGCACTGCAGCAGCAGCCNGCTTGTAGGTGACCGGAGAGAACCTCTCTCTCTCTCTCT 1362
Db 103528 CAGCACTGCAGCAGCAGCCNGCTTGTAGGTGACCGGAGAGAACCTCTCTCTCTCTCTCTCT 103587
QY 1363 CCCACTCTCTCTCCACT 1422
Db 103588 CCCACTCTCTCTCCACT 103647
QY 1423 AGATGAAAGAGGAGTCAAGTCTTTCAGTAGCCAAAACAAAACAAAACAAAACAAAAG 1482
Db 103648 AGATGAAAGAGGAGTCAAGTCTTTCAGTAGCCAAAACAAAACAAAACAAAACAAAAG 103707
QY 1483 CCGAATAAAGAAAAGAAAGATAATACTAGTCTTCTTATTTGCACTTCTTCTAGTGGACACT 1542

Db 103708	CGGAATAAAGAAAAAGATAAATCACTAGTCTTATTCCACCTACTTCAGTGGACACT	103767	Db 104782	ATTTCTGACAACGCCAAGGAGTTGTGTAGGCGAGTGTCCGTTGCCATGGGCGTJGGTGTG	104841
QY 1543	GAATTTGGAAGGTGGAGATTTTGTCTTTTAAAGATCTGGGCATCTTTTGAATCT	1602	QY 2623	GAGCGTTGGAGCATCTGAGTCCAGGGGAACAGCTTTCGGGGGATTCGATGTACGCCCA	2682
Db 103768	GAATTTGGAAGGTGGAGATTTTGTCTTTTAAAGATCTGGGCATCTTTGAATCT	103827	Db 104842	GAGCGTTGGAGCATCTGAGTCCAGGGGAACAGCTTTCGGGGGATTCGATGTACGCCCA	104901
QY 1603	ACCTTCAAGTATTAAAGAGACAGACTGTAGCTAGCAGGCGAGATCTTGTCCACGGTG	1662	QY 2683	CTTTTGGGAGCTTCACCCGCTGTGCGTCCCACTTCTTGTGCCCCATTTGGCCJAATGCAAA	2742
Db 103828	ACCTTCAAGTATTAAAGAGACAGACTGTAGCTAGCAGGCGAGATCTTGTCCACGGTG	103887	Db 104902	CTTTTGGGAGTTCACCCGCTGTGCGTCCCACTTCTTGTGCCCCATTTGGCCJAATGCAAA	104961
QY 1663	GTCTTCTTCGACGAGACTTTGAGGCTGTACAGCGCTTTTGTGCTGGTTCTCCCGCA	1722	QY 2743	GGTTCTCTGCTAGACGACCGCAGCAAGAGACTGTAGAGTACTGTCTGAGTATTCCTCT	2802
Db 103888	GTCTTCTTCGACGAGACTTTGAGGCTGTACAGCGCTTTTGTGCTGGTTCTCCCGCA	103947	Db 104962	GGTTCTCTGCTAGACGACCGCAGCAAGAGACTGTAGAGTACTGTCTGAGTATTCCTCT	105021
QY 1723	AGTTTCTTCTCTGGAGCTTCCCGCAGGTGGCAGCTAGCTCCAGGACTACCGCATCAT	1782	QY 2803	TTCAAGGGAGGTTTACACCAAGGGCTAGAGCGGAGAGCGCTAGGCTGTCTGGCAGCGCT	2862
Db 103948	AGTTTCTTCTCTGGAGCTTCCCGCAGGTGGCAGCTAGCTCCAGGACTACCGCATCAT	104007	Db 105022	TTCAAGGGAGGTTTACACCAAGGGCTAGAGCGGAGAGCGCTAGGCTGTCTGGCAGCGCT	105081
QY 1783	CACAGCCTGTGTAACCTTCTTGAGCAAGAGAGGGGAGCGGGTAAAGGAAGTAGGTGG	1842	QY 2863	GCAGCAGGAGCTTCCGGGACACTTGAACCTGCGCTTACCCCTGTCTCTCTCTCTCTCTCT	2922
Db 104008	CACAGCCTGTGTAACCTTCTTGAGCAAGAGAGGGGAGCGGGTAAAGGAAGTAGGTGG	104067	Db 105082	GCAGCAGGAGCTTCCGGGACACTTGAACCTGCGCTTACCCCTGTCTCTCTCTCTCTCTCT	105141
QY 1843	AAGATTCAGCAAGCTTCAAGGATGGAAGTGCAGTTAGGCTGGGAGAGGTCTACCTCGG	1902	QY 2923	GCACTGGAGCAGGAGCTGCGTACCAAGAGTCCGAGTCTACTACTACTACTACTACTACTACT	2982
Db 104068	AAGATTCAGCAAGCTTCAAGGATGGAAGTGCAGTTAGGCTGGGAGAGGTCTACCTCGG	104127	Db 105142	GCACTGGAGCAGGAGCTGCGTACCAAGAGTCCGAGTCTACTACTACTACTACTACTACTACT	105201
QY 1903	CCGCGCTTCAAGACCTTACCGAGAGCTTTCCAGAACTCTGTTCCAGAGCGTTCGCGAAGTG	1962	QY 2983	GCGGAGCCGCCGCCCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCG	3042
Db 104128	CCGCGCTTCAAGACCTTACCGAGAGCTTTCCAGAACTCTGTTCCAGAGCGTTCGCGAAGTG	104187	Db 105202	GCGGAGCCGCCGCCCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCG	105261
QY 1963	ATCCAGAACCCGGGCCAGGACCCAGAGCGCGGAGCGAGCAGTCCCGGGCCAGT	2022	QY 3043	CGCTTGGACTACGGCAGCGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCG	3102
Db 104188	ATCCAGAACCCGGGCCAGGACCCAGAGCGCGGAGCGAGCAGTCCCGGGCCAGT	104247	Db 105262	CGCTTGGACTACGGCAGCGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCG	105321
QY 2023	TTGCTGCTCTCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG	2082	QY 3103	GCGAGCCTGCATGCGCGGCTGCAGCGGACCCCGCTTCTGGGTACCCCTCAGCGCGCGCT	3162
Db 104248	TTGCTGCTGCT-----GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	104301	Db 105322	GCGAGCCTGCATGCGCGGCTGCAGCGGACCCCGCTTCTGGGTACCCCTCAGCGCGCGCT	105381
QY 2083	CAG	2142	QY 3163	TCCTCATCTTGGCAGCTCTTTCACAGCCGAGAGAGCCAGTGTGTATGACCGTGTGT	3222
Db 104302	CAG	104361	Db 105382	TCCTCATCTTGGCAGCTCTTTCACAGCCGAGAGAGCCAGTGTGTATGACCGTGTGT	105441
QY 2143	GCTGAGGATGTTTCTCCCAAGCCCATCGTAGAGGCCCCACAGGCTACCTTGTCTGGAT	2202	QY 3223	GCTGAGGATGTTTCTCCCAAGCCCATCGTAGAGGCCCCACAGGCTACCTTGTCTGGAT	105498
Db 104362	GCTGAGGATGTTTCTCCCAAGCCCATCGTAGAGGCCCCACAGGCTACCTTGTCTGGAT	104421	Db 105442	GCTGAGGATGTTTCTCCCAAGCCCATCGTAGAGGCCCCACAGGCTACCTTGTCTGGAT	105498
QY 2203	GAGGAACAGCAACCTTTCACAGCCGAGTCCGCGCTGAGTGCACCCCGAGAGAGTTGC	2262	QY 3283	GCGCGCGGAGGCGGAGCTGTAGCCCCCTAGGCTTACACTCGGCCCTCAGGGGCTG	3342
Db 104422	GAGGAACAGCAACCTTTCACAGCCGAGTCCGCGCTGAGTGCACCCCGAGAGAGTTGC	104481	Db 105499	GCGCGCGGAGGCGGAGCTGTAGCCCCCTAGGCTTACACTCGGCCCTCAGGGGCTG	105558
QY 2263	GTCCAGAGCCTGGAGCGCGCTGGCCGCCAGCAAGGGGCTGCCGAGCAGCTGCCAGCA	2322	QY 3343	GCGGGCCAGAAAGCGACTTTCACCGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3402
Db 104482	GTCCAGAGCCTGGAGCGCGCTGGCCGCCAGCAAGGGGCTGCCGAGCAGCTGCCAGCA	104541	Db 105559	GCGGGCCAGAAAGCGACTTTCACCGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	105618
QY 2323	CCTCCGAGCAGGATGACTCAGTGCCTTCCATCAGCTTGTCTGTGTGTGTGTGTGTGTGT	2382	QY 3403	AGAGTGCCTTATCCAGTCCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3462
Db 104542	CCTCCGAGCAGGATGACTCAGTGCCTTCCATCAGCTTGTCTGTGTGTGTGTGTGTGTGT	104601	Db 105619	AGAGTGCCTTATCCAGTCCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	105678
QY 2383	CCCGGCTTAAAGCAGTCTCCGCTGACCTTAAAGACATCTGAGCAGGCGCAGCAGCAGCAG	2442	QY 3463	TACTCCGGACCTTACGGGACATGCGGT 3490	
Db 104602	CCCGGCTTAAAGCAGTCTCCGCTGACCTTAAAGACATCTGAGCAGGCGCAGCAGCAGCAG	104661	Db 105679	TACTCCGGACCTTACGGGACATGCGGT 105706	
QY 2443	CAACTCTTCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	2502	RESULT 9		
Db 104662	CAACTCTTCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	104721	HUMARX		
QY 2503	AGGAGGCTTCGGGGCTCCACTTCCCAAGGACATTTACTTAGGGGCACTTTCGACC	2562	LOCUS		
Db 104722	AGGAGGCTTCGGGGCTCCACTTCCCAAGGACATTTACTTAGGGGCACTTTCGACC	104781	DEFINITION	Human androgen-receptor mRNA; complete cds.	
QY 2563	ATTCTGACAACGCCAAGGATTTGTAAAGCAGTGTCTGCTCCATGGCTGGGTGTG	2622	ACCESSION	M34233	
			VERSION	M34233.1	
			KEYWORDS	androgen receptor.	
			SOURCE	Human testis and prostate cancer cell line LNCaP, cdna to mRNA.	
			ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	

REFERENCE AUTHORS TITLE JOURNAL MEDLINE PUBMED	FEATURES source	gene	CDS	Location/Qualifiers	BASE COUNT	ORIGIN
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 3231) Govindan,M.V. Specific region in hormone binding domain is essential for hormone binding and trans-activation by human androgen receptor Mol. Endocrinol. 4 (3), 417-427 (1990) 90258935 2342476						
				Location/Qualifiers 1. .3231 /organism="Homo sapiens" /db_xref="taxon:9606" /map="Xq11.2-q12" 1. .3231 /gene="AR" 109..2829 /gene="AR" /note="androgen-receptor" /codon_start=1 /protein_id="AA51780.1" /db_xref="GI:179034" /db_xref="GDB:G00-120-556" /translation="MEVQLGLRVYPPSKTYRGAFQNLFSVREVIONPGRHPEA ASAPPPASLLQLQGGGQQQQQQQQQSTPQQQQQGGEDGSPQARRHPTGYLV LDEQQQSQPSALECHPERGPEGAASVASKGLPQOLPAPPPEDSDSAAPTSLTLL APTFPLGSSGADLLKDLILSASTMQLLQOOQEAEGSSGSSRGAREAGSAPTSSKNDY LGGTSTISDNAKELCKAVSVSMGLGVLEAHLSPGQLRGDCMVAPLLGVPAVPTP CAPLAECKGSLDDSAKSTDAEYSPFGKTYKLGEGESLGGSSRAAGSSGCTLEL PSTLSLYKSGALDDPAAYQRDYNFPLALAGPPPPPHPHARKLLENPLDYSAM AAAAOACRYGDLASLHGAAGGPGSGPSAAASSWHTLFTAEQGLYAGCGGGGGG GGGGGGGGEAGVAPYGYRPPQGLAQESDFTAPDVYPGMWSRVYPSPTCVKS EMGPMDWSGCPYGDMLRTARDHVLIDYFPQKTCITCGDEAGSGHYGALTCCGC KYVFRAGEQKVLICASRNDCTDKERKNCPCLRCYCEAGMTLGARKLKLGNPL KLOEGEASLSTPTEETQKLTVSHLEGYECOPILNVLAEPCVGCAGHNNQNP SFAALLSSMLGVRVHVYKNAKALPLGRNLHYDDQMAVIOYSHWGLMVRAMGRS FTNVSRLMYTAPDLVFNEXRBMKSRMYTSCVRRHRLSBEFQLQITPQBEFLCMKML LESIIIPVDGLNKNQFFDELBNMYIKELDRIIACKRKNPTSCRRFQITLKLDDSVQPI ARELHOFTFDLLIKSHMYSDVPEMMAEIIISVQPKILSKVPIYFHTQ"	713 a 941 c 883 g 694 t	
				Query Match 60.8%; Score 3089.4; DB 9; Length 3231; Best Local Similarity 97.7%; Pred. No. 0; Matches 3195; Conservative 0; Mismatches 21; Indels 53; Gaps 4;		
Qy	1756	ACGTAGCTGCAGCGACTACCGCATCATCAGCCTGTTGAACCTCTCTGACGACAGAG 1815				
Ddb	1	ACGTAGCTGCAGCGACTACCGCATCATCAGCCTGTTGAACCTCTCTGACGACAGAG 60				
Qy	1816	GGGAGGCGGGTAAGGGAATAGTGGAAGATTCACCAAGCTCAAGGATGAAGTGCAG 1875				
Ddb	61	GGGAGGCGGGTAAGGGAAGTAGTGGAAGATTCACCAAGCTCAAGGATGAAGTGCAG 120				
Qy	1876	TTAGGCTGGGAAGGCTACCTTCGGCCGCCCTCCAAAGACCTACCGAGAGCTTTCAG 1935				
Ddb	121	TTAGGCTGGGAAGGCTACCTTCGGCCGCCCTCCAAAGACCTACCGAGAGCTTTCAG 180				
Qy	1936	AATCTGTTCCAGAGCGTCCGGAAGTATCCAGAACCCGGGCCCCAGGACCCAGAGGCC 1995				
Ddb	181	AATCTGTTCCAGAGCGTCCGGAAGTATCCAGAACCCGGGCCCCAGGACCCAGAGGCC 240				
Qy	1996	GGCAGGCAGCACTCCCGGCCGAGTTTGTCTGCTGTCAGCAGCAGCAGCAGCAGCAG 2055				
Ddb	241	GGCAGGCAGCACTCCCGGCCGAGTTTGTCTGCTGTCAGCAGCAGCAGCAGCAGCAG 278				
Qy	2056	CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2115				
Ddb	279	-----GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 333				
Qy	2116	AGCCCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2175				
Ddb	334	AGCCCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 393				

D	b	1456	- - - - GGGGCGGGCGGGCGGGCGGGCGGGCGGGAGCGGGAGCTGTACGCCCCCTAC	1509
O	y	3316	GGCATCACTGCGCCCCCTCAGGGCTCGGGCGGAGAAAGCAATTACCCECACCTGAT	3375
D	b	1510	GGCATCACTGCGCCCCCTCAGGGCTCGGGCGGAGAAAGCAATTACCCECACCTGAT	1569
O	y	3376	GTGTGTACCCCTGGCGGCGATGGTGAGCAGAGTCCCCTATCCAGTCGCCACTGTGTCAA	3435
D	b	1570	GTGTGTGTACCCCTGGCGGCGATGGTGAGCAGAGTCCCCTATCCAGTCGCCACTGTGTCAA	1629
O	y	3436	AGCGAAATGGGCCCTCGATGATAGTTACTCGGACCTTAGGGGACATCGCTTTGGAG	3495
D	b	1630	AGCGAAATGGGCCCTCGATGATAGTTACTCGGACCTTAGGGGACATCGCTTTGGAG	1689
O	y	3496	ACTGCCAGGACCATGTTTTGCCAATGACTATTACTTTCACCCCGACAGAACCTGCCTG	3555
D	b	1690	ACTGCCAGGACCATGTTTTGCCAATGACTATTACTTTCACCCCGACAGAACCTGCCTG	1749
O	y	3556	ATCTGTGAGATGAAGCTTCTGGGTGTCACTATGGAGTCTCACATGTGGAAGCTGCAAG	3615
D	b	1750	ATCTGTGAGATGAAGCTTCTGGGTGTCACTATGGAGTCTCACATGTGGAAGCTGCAAG	1809
O	y	3616	GTCTCTTTCAAAGAGCGCTGAAGGAAACAAGAAGTACCTGTCGCCACACAGAAATGAT	3675
D	b	1810	GTCTCTTTCAAAGAGCGCTGAAGGAAACAAGAAGTACCTGTCGCCACACAGAAATGAT	1869
O	y	3676	TGCATAATTGATAAATTCGGAAGAAAAATGTCCAATTTGCTCTTCGGAATCTTAT	3735
D	b	1870	TGCATAATTGATAAATTCGGAAGAAAAATTTCCATCTTGTCTCTTCGGAATCTTAT	1929
O	y	3736	GAAGCAGGGATGACTCTGGGAGCCGGAGCTGARAACATTGGTAATCTGAACATPACAG	3795
D	b	1930	GAAGCAGGGATGACTCTGGGAGCCGGAGCTGARAACATTGGTAATCTGAACATPACAG	1989
O	y	3796	GAGGAGGAGAGGTTCCAGCACCCACCGCCACTGAGGAGACACCCAGAGCTGACA	3855
D	b	1990	GAGGAGGAGAGGTTCCAGCACCCACCGCCACTGAGGAGACACCCAGAGCTGACA	2049
O	y	3856	GTGTTCACACATTGAAGSGCTAATAATGTGAGCCCATCTTCTGTAATGTCTTGGAGCCATT	3915
D	b	2050	GTGTTCACACATTGAAGSGCTAATAATGTGAGCCCATCTTCTGTAATGTCTTGGAGCCATT	2109
O	y	3916	GAGCCAGGTGTGTGTCTGGACACACAAACAGCCGACCTTTCGAGCCATT	3975
D	b	2110	GAGCCAGGTGTGTGTCTGGACACACAAACAGCCGACCTTTCGAGCCATT	2169
O	y	3976	CTCTCTAGCTCAATGAACCTGGGAGAGACAGCTGTGTACAGTGGTCAAGTGGCCAAAG	4035
D	b	2170	CTCTCTAGCTCAATGAACCTGGGAGAGACAGCTGTGTACAGTGGTCAAGTGGCCAAAG	2229
O	y	4036	GCCTTGCTTGGCTTCGCAACTTACAGTGGAGACACCATGCTGTCAATGTCTCACTCTCC	4095
D	b	2230	GCCTTGCTTGGCTTCGCAACTTACAGTGGAGACACCATGCTGTCAATGTCTCACTCTCC	2289
O	y	4096	TGATGGGCTCATGGTGTTCGCATGGGTGGCGATCTTCAACAATGTCACCTCCAGG	4155
D	b	2290	TGATGGGCTCATGGTGTTCGCATGGGTGGCGATCTTCAACAATGTCACCTCCAGG	2349
O	y	4156	ATGCTCTACTTCGCCCTTGATCTGGTTTTCAATGAGTACCGCATGCACAAGTCCCGATG	4215
D	b	2350	ATGCTCTACTTCGCCCTTGATCTGGTTTTCAATGAGTACCGCATGCACAAGTCCCGATG	2409
O	y	4216	TACAGCCAGTGTGTCCGAATGAGCAGCTCTCTCAGAGTTTGGATGGCTCCAAATCAOC	4275
D	b	2410	TACAGCCAGTGTGTCCGAATGAGCAGCTCTCTCAGAGTTTGGATGGCTCCAAATCAOC	2469
O	y	4276	CCCCAGGAATTCCTGTGATGAAAGCACTGCTACTCTTCAGCATTATTCAGTGGATGG	4335
D	b	2470	CCCCAGGAATTCCTGTGATGAAAGCACTGCTACTCTTCAGCATTATTCAGTGGATGG	2529
O	y	4336	CTGAAAATCAAAAATCTTTGTATGAACTTCGAATGAACATCAAGGAACCTCGATCGT	4395

Db	2530	CTGAAAAATCAAAAATCTTTTGATGAACCTTCCGTAATGAACATACATCAAGGNACTGATCGT	2589
Qy	4396	ATCATTTGCATCAAAAGAAAAATCCCAATCCTGCTCAAGACGCTTCTACCAAGCTCACC	4455
Db	2590	ATCATTTGCATCAAAAGAAAAATCCCAATCCTGCTCAAGACGCTTCTACCAAGCTCACC	2649
Qy	4456	AAGCTCTGGACTCCGTCGACGCTTATTCGAGAGAGCTGCATCAGTTTCACTTTTGACCTG	4515
Db	2650	AAGCTCTGGACTCCGTCGACGCTTATTCGAGAGAGCTGCATCAGTTTCACTTTTGACCTG	2709
Qy	4516	CTAATCAAGTCACACATGTGAGCTTCCCGAAATGATGGCAGAGATCATCTCT	4575
Db	2710	CTAATCAAGTCACACATGTGAGCTTCCCGAAATGATGGCAGAGATCATCTCT	2769
Qy	4576	GTGCAAGTGCCCAAGATCCCTTTCTGGGAAAGTCAAGCCCATCTATTTCCACACCCAGTGA	4635
Db	2770	GTGCAAGTGCCCAAGATCCCTTTCTGGGAAAGTCAAGCCCATCTATTTCCACACCCAGTGA	2829
Qy	4636	AGCATTGGAACCCCTATTTCCCCACCCAGCTCATGCCCTTTTCAGATGTCTTTCGCT	4695
Db	2830	AGCATTGGAACCCCTATTTCCCCACCCAGCTCATGCCCTTTTCAGATGTCTTTCGCT	2889
Qy	4696	GTATTAACCTGCACTACCTCTGCAGTGCCCTTGGGAAATTTCCCTCTATTTGAATGACAG	4755
Db	2890	GTATTAACCTGCACTACCTCTGCAGTGCCCTTGGGAAATTTCCCTCTATTTGAATGACAG	2949
Qy	4756	TCGTGTCATGAACATGTCCTGAATCTATTTGCTGGGCTTTTTCCTCTCTCTCTCTCT	4815
Db	2950	TCGTGTCATGAACATGTCCTGAATCTATTTGCTGGGCTTTTTCCTCTCTCTCTCTCT	3009
Qy	4816	TTCTTTTCT	4875
Db	3010	TTCTTTTCT	3069
Qy	4876	ATTGTGGCTCCATCTGCTGTTTGAATGTGTGTATGCTTTAAATCTGTGATGATCCT	4935
Db	3070	ATTGTGGCTCCATCTGCTGTTTGAATGTGTGTATGCTTTAAATCTGTGATGATCCT	3128
Qy	4936	CATATGGCCCACTGTCAGTGTGCTTGTGTTTACACACACTACCTCTCTGCGCCACACACAA	4995
Db	3129	CATATGGCCCACTGTCAGTGTGCTTGTGTTTACACACACTACCTCTCTGCGCCACACAA	3187
Qy	4996	CGTTTACTATCTTATGTCACGGGAAGTT	5024
Db	3188	CGTTTACTATCTTATGTCACGGGAAGTT	3216
RESULT 10			
HUMANRE			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
FEATURES			
source			
gene			

Db	1708	ATCTGTGGAGATGAAGCTTCTGGGTGTCACTATGAGCTCTCACATGTGGAGCTGCAAG	1767
QY	3616	GTCTTCTTCAAAAGAGCGCTGAAGGGAAACAGAAAGTACCTGTGCGCCACGAGAAATGAT	3675
Db	1768	GTCTTCTTCAAAAGAGCGCTTGAAGGGAAACAGAAAGTACCTGTGCGCCACGAGAAATGAT	1827
QY	3676	TGCACATTTGATAAATTCGGAAGGAAAATTTGCCATCTTGTCTCTTCGGAATGTTAT	3735
Db	1828	TGCACATTTGATAAATTCGGAAGGAAAATTTGCCATCTTGTCTCTTCGGAATGTTAT	1887
QY	3736	GAAGCAGGGAATGACTCTGTGGAGCCGGAAGCTGAAGAACTTGGTAATCTGAAACTACAG	3795
Db	1888	GAAGCAGGGAATGACTCTGTGGAGCCGGAAGCTGAAGAACTTGGTAATCTGAAACTACAG	1947
QY	3796	GAGGAAGAGAGGCTTCCAGCACACAGACGCCCACTGAGGAGACAACCCAGAAGCTGACA	3855
Db	1948	GAGGAAGAGAGGCTTCCAGCACACAGACGCCCACTGAGGAGACAACCCAGAAGCTGACA	2007
QY	3856	GTGTTCACACATTTGAAGGCTATGAATGTCAAGCCCATCTTCTGAATGTCTTGAAGCCATT	3915
Db	2008	GTGTTCACACATTTGAAGGCTATGAATGTCAAGCCCATCTTCTGAATGTCTTGAAGCCATT	2067
QY	3916	GAGCCAGGTGTAGTGTGTGTGGACAGCAACAACACAGCCGAGCTCTTTGACGCTTGG	3975
Db	2068	GAGCCAGGTGTAGTGTGTGTGGACAGCAACAACACAGCCGAGCTCTTTGACGCTTGG	2127
QY	3976	CTCTCTAGCCCTCAATGAACGTGGAGAGACAGACCTTGTACAGTGGTCAAGTGGGCCAAG	4035
Db	2128	CTCTCTAGCCCTCAATGAACGTGGAGAGACAGACCTTGTACAGTGGTCAAGTGGGCCAAG	2187
QY	4036	GCCTTGCCCTGGCTTCCGCAACTTTACAGCTGGAGACAGACAGTGGCTGTCTCATTCAGTACTCC	4095
Db	2188	GCCTTGCCCTGGCTTCCGCAACTTTACAGCTGGAGACAGACAGTGGCTGTCTCATTCAGTACTCC	2247
QY	4096	TGATGGGCTCATGTGTTTGCCATGGCTGGGATCCTTCAACAAATGTCAACTTCCAGG	4155
Db	2248	TGATGGGCTCATGTGTTTGCCATGGCTGGGATCCTTCAACAAATGTCAACTTCCAGG	2307
QY	4156	ATGCTCTACTTCGCCCTGATCTGGTTTTCATAGTAGTACCGCATGCCAAGTCCCAATCACC	4215
Db	2308	ATGCTCTACTTCGCCCTGATCTGGTTTTCATAGTAGTACCGCATGCCAAGTCCCAATCACC	2367
QY	4216	TACAGCCAGTGTTCGGATGAGGCACCTCTCTCAAGAGTTTGGATGGCTCCCAATCACC	4275
Db	2368	TACAGCCAGTGTTCGGATGAGGCACCTCTCTCAAGAGTTTGGATGGCTCCCAATCACC	2427
QY	4276	CCCCAGGAATTCCTGTGCATGAAGACACTGTGA-----CTCTTC	4314
Db	2428	CCCCAGGAATTCCTGTGCATGAAGACACTGTGA-----CTCTTC	2487
QY	4315	AGCATTTATCCAGTGGATGGCTGAAATAATCAAAATTCCTTTGATGAACCTCGAATGAAC	4374
Db	2488	AGCATTTATCCAGTGGATGGCTGAAATAATCAAAATTCCTTTGATGAACCTCGAATGAAC	2547
QY	4375	TACATCAAGGAACTCGATCGTATCATTTGCATGCAAAAGAAAAATCCCAATCTCGTCA	4434
Db	2548	TACATCAAGGAACTCGATCGTATCATTTGCATGCAAAAGAAAAATCCCAATCTCGTCA	2607
QY	4435	AGAGCCTTCTACCAAGCTCACCAAGCTCCTGAGCTCCGTGAGGCTATTCGGAGAGAGCTG	4494
Db	2608	AGAGCCTTCTACCAAGCTCACCAAGCTCCTGAGCTCCGTGAGGCTATTCGGAGAGAGCTG	2667
QY	4495	CATCAGTTTCACTTTTGGACCTGCTAATCAAGTTCACACATGTGAGCGTGGACTTTCGGAA	4554
Db	2668	CATCAGTTTCACTTTTGGACCTGCTAATCAAGTTCACACATGTGAGCGTGGACTTTCGGAA	2727
QY	4555	ATGATGGCAGAGATCATCTCTGTGCAAGTGCCTCAAGATCCTTTCTTGGGAAAGTCAAGCCC	4614
Db	2728	ATGATGGCAGAGATCATCTCTGTGCAAGTGCCTCAAGATCCTTTCTTGGGAAAGTCAAGCCC	2787
QY	4615	ATCTATTTTCCACCCAGTGAAGCATTTGGAACCCCTATTTTCCCCACCCAGCTCATGCC	4674
Db	2788	ATCTATTTTCCACCCAGTGAAGCATTTGGAACCCCTATTTTCCCCACCCAGCTCATGCC	2847

[illegible]

Db	2125		AAATGGGCGAAGCGCTTGCTGGCTTCGCAACTTACACGTGGACGACGATGGCTGTC	2184
QY	4084		ATTACAGTACTCTCGATGGGCTCATGGTGTTCCTCCATGGGCTGGCGATCCCTACCAAT	4143
Db	2185		ATTACAGTACTCTCGATGGGCTCATGGTGTTCCTCCATGGGCTGGCGATCCCTACCAAT	2244
QY	4144		GTCACACTCCAGATGCTCTACTTCGCCCTCTGATCTGGTTTCAATGAGTACCGCATGCAC	4203
Db	2245		GTCACACTCCAGATGCTCTACTTCGCCCTCTGATCTGGTTTCAATGAGTACCGCATGCAC	2304
QY	4204		AAATCCCGGATGTACAGCAGTGTCCGAAATGCTCTCTCTCAAGAGTTTGGATGG	4263
Db	2305		AAATCCCGGATGTACAGCAGTGTCTCCGAAATGCTCTCTCTCAAGAGTTTGGATGG	2364
QY	4264		CTCCAAATCACCCTCCAGGAATTCCTGTGCATGAAAGCACTGCTACTCTTCAGCATTTAT	4323
Db	2365		CTCCAAATCACCCTCCAGGAATTCCTGTGCATGAAAGCACTGCTACTCTTCAGCATTTAT	2424
QY	4324		CCAGTGGATGGCTGAAAAATCAAAATCTTTGATGAATTCGAAATGAATACATCAAG	4383
Db	2425		CCAGTGGATGGCTGAAAAATCAAAATCTTTGATGAATTCGAAATGAATACATCAAG	2484
QY	4384		GAATCTCATGTATCATTCGATGCAAAAGAAAAATCCACATCTCTGCTCAAGACGCTTC	4443
Db	2485		GAATCTCATGTATCATTCGATGCAAAAGAAAAATCCACATCTCTGCTCAAGACGCTTC	2544
QY	4444		TACAGCTCACCAAGCTCTCGACTCCGTCAGGCTATTCGAGAGAGCTGCATCAGTTC	4503
Db	2545		TACAGCTCACCAAGCTCTCGACTCCGTCAGGCTATTCGAGAGAGCTGCATCAGTTC	2604
QY	4504		ACTTTTCACCTGTATCAAGTCACACATGCTGAGCGTGGACTTTCGGAAATGATGGCA	4563
Db	2605		ACTTTTCACCTGTATCAAGTCACACATGCTGAGCGTGGACTTTCGGAAATGATGGCA	2664
QY	4564		GAGATCATCTCTGCAAGTGGCCCAAGATCCCTTTCTGGAAAGTCAAGCCATCTATTTC	4623
Db	2665		GAGATCATCTCTGCAAGTGGCCCAAGATCCCTTTCTGGAAAGTCAAGCCATCTATTTC	2724
QY	4624		CACACCCAGTGAAGCATTTGAAA-CCCTATTTCGCCACCCAGCTCATGCCCTTTTCAG	4682
Db	2725		CACACCCAGTGAAGCATTTGAAAACCCCTATTTCGCCACCCAGCTCATGCCCTTTTCAG	2784
QY	4683		ATCTCTTCGCTGTATTAACCTCTGCATCTCTCTGCTGAGTGCCTTGGGGAATTCCTTC	4742
Db	2785		ATCTCTTCGCTGTATTAACCTCTGCATCTCTCTGAGTGCCTTGGGGAATTCCTTC	2844
QY	4743		TATTGATGTACAGTCTGTCATGAAC	4767
Db	2845		TATTGATGTACAGTCTGTCATGAAC	2869
RESULT	13			
REFERENCE	AF162704			
LOCUS	AF162704	2827 bp	mrna	linear
DEFINITION	Homo sapiens cell-line MDA-MB-453 androgen receptor mRNA, complete cds.			
ACCESSION	AF162704			
VERSION	AF162704.1	GI:5639998		
KEYWORDS				
SOURCE	Homo sapiens			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
TITLE	1 (bases 1 to 2827)			
DESCRIPTION	Jin, C.H., Urcan-Bisel, M.S. and Schrader, W.T.			
KEYWORDS	Androgen receptor sequences in human mammary carcinoma MDA-MB-453 cells			
REFERENCE	Unpublished			
AUTHORS	2 (bases 1 to 2827)			
TITLE	Jin, C.H., Urcan-Bisel, M.S. and Schrader, W.T.			
DESCRIPTION	Direct Submission			
KEYWORDS	Submitted (25-JUN-1999) Endocrine Research, Ligand Pharmaceuticals			

QY	2401	TCGGTGAACCTTAAGACATCTTGAGGAGCCAGCACATGCAACTCCTTCAGCAACAG	2460
Db	514	TCGGTGAACCTTAAGACATCTTGAGGAGCCAGCACATGCAACTCCTTCAGCAACAG	573
QY	2461	CAGCAGGACGATATCCGAGGACGACGACGCGGAGAGCGAGGAGCCTCGGGGCT	2520
Db	574	CAGCAGGACGATATCCGAGGACGACGACGCGGAGAGCGAGGAGCCTCGGGGCT	633
QY	2521	CCACTTCCTCAAGACATTAATCTAGGGGACCTTCGACCATTTCTGACACGCCAAG	2580
Db	634	CCACTTCCTCAAGACATTAATCTAGGGGACCTTCGACCATTTCTGACACGCCAAG	693
QY	2581	GAGTGTGTAAAGCAGTCTCGGTCTCCATGGGCTGGGTGTGGAGGCGTTGGAGCATCTG	2640
Db	694	GAGTGTGTAAAGCAGTCTCGGTCTCCATGGGCTGGGTGTGGAGGCGTTGGAGCATCTG	753
QY	2641	AGTCCAGGGGAACAGCTTCGGGGGGATGTCATACGCCCCCACTTTTGGGAGTTCCACCC	2700
Db	754	AGTCCAGGGGAACAGCTTCGGGGGGATGTCATGTCGCCCACTTTTGGGAGTTCCACCC	813
QY	2701	GCTGTGCTGCCACCTCTTGTCGCCCATTTGGCCGAATGCAAAAGTTCTGCTAGACGAC	2760
Db	814	GCTGTGCTGCCACCTCTTGTCGCCCATTTGGCCGAATGCAAAAGTTCTGCTAGACGAC	873
QY	2761	AGCGAGGCAAGACGACTGAGATCTCTCAGTATTTCCCTTTCAAGGAGGTTACACC	2820
Db	874	AGCGAGGCAAGACGACTGAGATCTCTCAGTATTTCCCTTTCAAGGAGGTTACACC	933
QY	2821	AAAGGGCTAGAAGCGAGAGCCTAGGCTGCTCTGGCAGCGCTCGACAGGGAGTCCGGG	2880
Db	934	AAAGGGCTAGAAGCGAGAGCCTAGGCTGCTCTGGCAGCGCTCGACAGGGAGTCCGGG	993
QY	2881	ACACTTGAATGCGCTTACCTGTCTCTACAGTCCGAGACCTTGACACGAGCAGCT	2940
Db	994	ACACTTGAATGCGCTTACCTGTCTCTACAGTCCGAGACCTTGACACGAGCAGCT	1053
QY	2941	CGGTACCAGAGTCGCGACTACTACAATTTCCACTTGGCTCTGGCGGACCGCGCCCT	3000
Db	1054	CGGTACCAGAGTCGCGACTACTACAATTTCCACTTGGCTCTGGCGGACCGCGCCCT	1113
QY	3001	CGCGCGCTCCCATCCCCACGCTCGCATCAGCTGGAGAACCGCTTGACATACGCGACG	3060
Db	1114	CGCGCGCTCCCATCCCCACGCTCGCATCAGCTGGAGAACCGCTTGACATACGCGACG	1173
QY	3061	GCTTGGCGCTCGCGCGCGCAGTCCGCTATGGGACCTTGGCGACCTGATGCGCGG	3120
Db	1174	GCTTGGCGCTCGCGCGCGCAGTCCGCTATGGGACCTTGGCGACCTGATGCGCGG	1233
QY	3121	GGTGAGGGGACCGGTTCTGGGTCACTCAGCGCGGCTTCTCATCTGGCACACT	3180
Db	1234	GGTGAGGGGACCGGTTCTGGGTCACTCAGCGCGGCTTCTCATCTGGCACACT	1293
QY	3181	CTCTTTCACAGCGGAAGAGCCAGTTGTATGACCGCTGTGTGTGGGGTGTGTGGC	3240
Db	1294	CTCTTTCACAGCGGAAGAGCCAGTTGTATGACCGCTGTGTGTGGGGTGTGTGGC	1350
QY	3241	GGCGGGCGGGA	3300
Db	1351	GGCGGGCGGGA	1389
QY	3301	GCTGTAGCCCCCTACGCTACACTTCGCCCCCTCTAGGGGCTGGCGGCGAGGAAGCGAC	3360
Db	1390	GCTGTAGCCCCCTACGCTACACTTCGCCCCCTCTAGGGGCTGGCGGCGAGGAAGCGAC	1449
QY	3361	TTACCGCGACCTGATGTGTGTACCTTGGGGCATGGTCAGCAGAGTGCCTATCCAGT	3420
Db	1450	TTACCGCGACCTGATGTGTGTACCTTGGGGCATGGTCAGCAGAGTGCCTATCCAGT	1509
QY	3421	CCCACTTGTGTAAAAGCGAAATGGGCCCTTGGATGGATAGCTACTCCGGACCTTACGGG	3480
Db	1510	CCCACTTGTGTAAAAGCGAAATGGGCCCTTGGATGGATAGCTACTCCGGACCTTACGGG	1569
QY	3481	GACATGCGTTTGAGACTGCCAGGACCATTTTGGCCATTGACTATTACTTTCCACCC	3540

Db	1570	GACATGGCTTGGAGACTGCCAGGACCATGTTTGTGCCATTGACTATTACTTTTCAACC	1629
Qy	3541	CAGAAGACCTCGCTGATCTGTGGAGATGAAGCTTCTGGGTGTCACATATGAGAGCTCTACA	3600
Db	1630	CAGAAGACCTGCCCTGATCTGTGGAGATGAAGCTTCTGGGTGTCACATATGAGACTCTACA	1689
Qy	3601	TGTGGAAGCTCAAGGTCTTCTTCAAAGAGCCGCTGAAGGGAACAGAAAGTACCTGTGC	3660
Db	1690	TGTGGAAGCTGCAAGGTCTTCTTCAAAGAGCCGCTGAAGGGAACAGAAAGTACCTGTGC	1749
Qy	3661	GCCAGCAGAAATGATTGGACTATTGATTAATTCGGAAGGAAAAATTTGCCATCTTCTCGT	3720
Db	1750	GCCAGCAGAAATGATTGGACTATTGATTAATTCGGAAGGAAAAATTTGCCATCTTCTCGT	1809
Qy	3721	CTTTCGGAATGTTATGAAGCAGGATGACTCTGGAGCCCGGAAGCTGAAGAAACTTGT	3780
Db	1810	CTTTCGGAATGTTATGAAGCAGGATGACTCTGGAGCCCGGAAGCTGAAGAAACTTGT	1869
Qy	3781	AATCTGAAACTACAGAGGAGGAGAGCTTCCAGCACCACAGCCCACTGAGGAGACA	3840
Db	1870	AATCTGAAACTACAGAGGAGGAGAGGCTTCCAGCACCACCAGCCCACTGAGGAGACA	1929
Qy	3841	ACCCAGAAGCTGACAGTGTCAACATTTGAAGGCTATGAATGTAGCCCATCTTCTGAAT	3900
Db	1930	ACCCAGAAGCTGACAGTGTCAACATTTGAAGGCTATGAATGTAGCCCATCTTCTGAAT	1989
Qy	3901	GTCTCTGGAAGCCATTGAGCCAGCTGATGTGTCTGACACGACACAAACACAGCCCGAC	3960
Db	1990	GTCTCTGGAAGCCATTGAGCCAGCTGATGTGTCTGACACGACACAAACACAGCCCGAC	2049
Qy	3961	TCCTTTGCAGCCTTGCTCTCTAGCCTCAATGAACCTGGGAGAGACAGCTTGTACAGTG	4020
Db	2050	TCCTTTGCAGCCTTGCTCTCTAGCCTCAATGAACCTGGGAGAGACAGCTTGTACAGTG	2109
Qy	4021	GTCAAGTGGCCAAAGCCCTTGCTGTGCTTCCCAACTTACAGTGGAGAGACAGATGGCT	4080
Db	2110	GTCAAGTGGCCAAAGCCCTTGCTGTGCTTCCCAACTTACAGTGGAGAGACAGATGGCT	2169
Qy	4081	GTCATTCAGTACTCCTCTGGATGGGGCTCATGGTGTTCATGGCTGGGATGCCCTCAC	4140
Db	2170	GTCATTCAGTACTCCTCTGGATGGGGCTCATGGTGTTCATGGCTGGGATGCCCTCAC	2229
Qy	4141	AATGTCAACTCCAGGATGCTACTTCCGCCCTGATCTCGTGTTCATAGAGTACCGCATG	4200
Db	2230	AATGTCAACTCCAGGATGCTACTTCCGCCCTGATCTCGTGTTCATAGAGTACCGCATG	2289
Qy	4201	CACAAGTCCCGGATGTACAGCAGTGTGCCGAATGAGGCACCTCTCTCAAGATTTGGA	4260
Db	2290	CACAAGTCCCGGATGTACAGCAGTGTGCCGAATGAGGCACCTCTCTCAAGATTTGGA	2349
Qy	4261	TGGCTCCAAATCACCCCCAGGAATTCCTGTGCATGAAAGCACTGCTACTCTTACGATT	4320
Db	2350	TGGCTCCAAATCACCCCCAGGAATTCCTGTGCATGAAAGCACTGCTACTCTTACGATT	2409
Qy	4321	ATTCCAGTGGATGGGCTGAAAAATTCAAAATTCCTTTGATGAATTCGAATGAATACATC	4380
Db	2410	ATTCCAGTGGATGGGCTGAAAAATTCAAAATTCCTTTGATGAATTCGAATGAATACATC	2469
Qy	4381	TAGGAACCTCGATCGTATCATTTGCATGCAAAAGAAAAATCCCATCTCTCTCAAGCCG	4440
Db	2470	TAGGAACCTCGATCGTATCATTTGCATGCAAAAGAAAAATCCCATCTCTCTCAAGCCG	2529
Qy	4441	TTCTACAGCTACCAAGCTCCTGGACTCCGTGCAGCCTATTCGGAGAGAGCTGCATCAG	4500
Db	2530	TTCTACAGCTACCAAGCTCCTGGACTCCGTGCAGCCTATTCGGAGAGAGCTGCATCAG	2589
Qy	4501	TTTCACTTTTGACTGCTTAATCAAGTACACATGTTGAGCGGTGGACTTTCCGGAATGATG	4560
Db	2590	TTTCACTTTTGACTGCTTAATCAAGTACACATGTTGAGCGGTGGACTTTCCGGAATGATG	2649
Qy	4561	GCAGAGATCATCTCTGTGCAAGTGCCCAAGTCTTCTCTGGGAAGTCAACCCCATCTAT	4620


```
Db 2650 GCAGAGATCATCTCTGTGCAAGTCCCAAGATCCTTTCTGGGAAAGTCAAGCCCATCTAT 2709
QY 4621 TTCCACACCCAGTGAAGATGGAAACCCATTTTCCACCCAGTTCATGCCCCCTTTC 4680
Db 2710 TTCCACACCCAGTGAAGATGGAAACCCATTTTCCACCCAGTTCATGCCCCCTTTC 2769
QY 4681 AGATGCTCTTCTGCTGTATTAACCTCTGCATCTACTCTCTGCAAGTGGGGAATTT 4738
Db 2770 AGATGCTCTTCTGCTGTATTAACCTCTGCATCTACTCTCTGCAAGTGGGGAATTT 2827

RESULT 14
LOCUS MFU941179
DEFINITION Macaca fascicularis androgen receptor mRNA, complete cds.
ACCESSION U941179
VERSION U941179.1 GI:3861480
KEYWORDS
SOURCE Macaca fascicularis.
ORGANISM Macaca fascicularis.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
Cercopithecinæ; Macaca.
1 (bases 1 to 2821)
Choong,C.S., Kempainen,J.A. and Willson,E.M.
Evolution of the primate androgen receptor: a structural basis for
disease
J. Mol. Evol. 47 (3), 334-342 (1998)
98404153
MEDLINE 9732460
PUBMED
REFERENCE 2 (bases 1 to 2821)
Choong,C.S., Kempainen,J.A. and Willson,E.M.
Direct Submission
Submitted (18-MAR-1997) Laboratories for Reproductive Biology,
University of North Carolina at Chapel Hill, CB 7500, MSRB, Rm 370,
UNC-CH, Chapel Hill, NC 27599, USA
LOCATION/Qualifiers
1..2821
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
1..2688
/codon_start=1
/product="androgen receptor"
/protein_id="AAC73050.1"
/db_xref="GI:3861481"
/translation="MEVQLGLGRVYPRPSKTYKAFONLQSVREIVQNPRPHEA
ASAPFGSLQQQQOQOQETSPQQOQOQOQEDSGSPQHRRGPTGYLVLDEEQSPQ
QSAPECHPERGCVPEPRAVAGKLPQLFAPDEDDSAAPSTLSLLGTFPLSSC
STDLDLISEATMOLLOOQOQEAQVSESSGSSGAREASGAPTSSKDNVLTSTISD
AKELCKAVSYMGLVLEHLSLPGELRGDCMYAPVLPVAVRPTCPAPLAECKGS
LDDASNGKSTEDTAYSPFKGYTKLGESLGSAAAGSGTLELPTLSLYKSG
ALDEAAVOSRDYNYFPALAGPPPPPPHARIKLENLDYGGSSAWAAAAQCRIG
DASLHGAGAAGPGSPSAAASSWHTLFAERGLYPCGGGGGGGGGGGAGEA
GAVAPYTRPPQGLAGQGTAPDVWYPGMYRVPYSPGTCVSENGPMDYSYG
PYGDMRLTARDHVLPIDYFPPOKTCILGDEASGCHYALTCGSKVFPKRAECK
QKYLCAANDCTIDKFRKNCPSLRCYFAGMTLGLKRLKLNGLKLEGEFASST
LGERQLVHWKALPGFNRHLDVDDQAVIOYSWGLMYFAMGWSFTNVSRLYF
APDLVFNIRHKSRLMYSQVVRHLSQEFGLQITPOEFLCMKALLLESIPVDGLK
NOKFDELRMYIKELDRITIAKRNKPTSCSRFYLTKLLDSVQPIARELQIFEDL
LIKHSVMSVDFPENMAELISVQPKLSGKVKPIYHTQ"
BASE COUNT 623 a 829 c 799 g 570 t
ORIGIN

Query Match
Best Local Similarity 51.4%; Score 2612.4; DB 9; Length 2821;
Matches 2771; Conservative 0; Mismatches 46; Indels 91; Gaps 4;

QY 1864. ATGGAGAGTCACTAGGCTGGGAGGCTACCTCCGCGCGCTCCAGACCTACCGA 1923
Db 1 ATGGAGGTGCATAGGCTGGGAGGCTACCTCCGCGCGCGCTCCAGACCTACCGA 1923
QY 1924 GGAGCTTCCAGATCTCTTCCAGAGGTGGGAGTATCCAGAACCCGCCCCAGG 1983
```

Fri May 9 14:20:38 2003

Db 1081 CCACGCCCTCCCATCCACACGCTCCGATCAAGCTGGAGAACCCCGCTGGACTATGGCAGC 1140
 Qy 3061 GCTTGGGGGCTGGCGGCGCAGTCCCGCTATGGGACCTGGGAGCCTGCATGGCGGG 3120
 Db 1141 GCTTGGGGGCTGGCGGCGCAGTCCCGCTATGGGACCTGGGAGCCTGCATGGCGGG 1200
 Qy 3121 GGTGCGAGCGGACCCGGTCTGGGTACCTCAGCCGCGCTTCCTCATCTGCGACACT 3180
 Db 1201 GGTGCGAGCGGACCCGGTCTGGGTACCTCAGCCGCGCTTCCTCATCTGCGACACT 1260
 Qy 3181 CTCTTACAGCCCAAGAGCGGCTGTATGGACCGCTGTGGTGGTGGTGGTGGTGG 3240
 Db 1261 CTCTTACAGCCCAAGAGCGGCTGTATGGACCGCTGTGGTGGTGGTGGTGGTGG 1305
 Qy 3241 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGGA 3300
 Db 1306 -----GGGCGGCGGCGGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGGA 1353
 Qy 3301 GCTGTAGCCCTACGCTACACTCGGCCCTCTCAGGGGCTGGCGGCGGCGGCGGCGG 3360
 Db 1354 GCTGTAGCCCTACGCTACACTCGGCCCTCTCAGGGGCTGGCGGCGGCGGCGGCGG 1413
 Qy 3361 TTCACCGCACCTGTATGGTACCTTGGCGGCTGCTGAGCAGAGTGGCTATCCAGT 3420
 Db 1414 TTCACCGCACCTGTATGGTACCTTGGCGGCTGCTGAGCAGAGTGGCTATCCAGT 1473
 Qy 3421 CCCACTTGTCAAAGCGAATGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3480
 Db 1474 CCCACTTGTCAAAGCGAATGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1533
 Qy 3481 GACATGCGCTTGGAGCTGCGCAGGACCATGTTTGGCCATGACTATTTTCCACCC 3540
 Db 1534 GACATGCGCTTGGAGCTGCGCAGGACCATGTTTGGCCATGACTATTTTCCACCC 1593
 Qy 3541 CAGAGACCTGCTGATCTGTGAGATGAAGCTTCTGGGCTGCTACTATGAGCTCTACA 3600
 Db 1594 CAGAGACCTGCTGATCTGTGAGATGAAGCTTCTGGGCTGCTACTATGAGCTCTACA 1653
 Qy 3601 TGTGGAAGCTGCAAGCTTCTCAAAGAGCGGCTGAAGGGAACAGAGTACTCTGTC 3660
 Db 1654 TGTGGAAGCTGCAAGCTTCTCAAAGAGCGGCTGAAGGGAACAGAGTACTCTGTC 1713
 Qy 3661 GCACGAGAAATGATGCACTATTGATAAATTCGAAGGAAATTTGTCATCTTTCGT 3720
 Db 1714 GCACGAGAAATGATGCACTATTGATAAATTCGAAGGAAATTTGTCATCTTTCGT 1773
 Qy 3721 CTTGGAATGTTATGAAGCAGGATGACTCTGGGAGCCCGGAGCTGAAGAAATTTGCT 3780
 Db 1774 CTTGGAATGTTATGAAGCAGGATGACTCTGGGAGCCCGGAGCTGAAGAAATTTGCT 1833
 Qy 3781 AATCTGAACCTACAGGAGGAGAGGCTTCCAGCACCAGCCCGCTGAGGAGACA 3840
 Db 1834 AATCTGAACCTACAGGAGGAGAGGCTTCCAGCACCAGCCCGCTGAGGAGACA 1893
 Qy 3841 ACCCAGAGCTGACAGTGTACACATGAGGCTATGATGTCAGCCCATCTTTCTGAAT 3900
 Db 1894 GCCCAGAGCTGACAGTGTACACATGAGGCTATGATGTCAGCCCATCTTTCTGAAT 1953
 Qy 3901 GTCCTGGAAGCCATTGAGCCAGGCTAGTGTGTGCTGGACACCAACACAGCCCGAC 3960
 Db 1954 GTCCTGGAAGCCATTGAGCCAGGCTAGTGTGTGCTGGACACCAACACAGCCCGAC 2013
 Qy 3961 TCCCTTGGAGCCTGCTCTAGCCTCAATGAACTGGGAGAGACAGCTTGTACAGCTG 4020
 Db 2014 TCCCTTGGAGCCTGCTCTAGCCTCAATGAACTGGGAGAGACAGCTTGTACAGCTG 2073
 Qy 4021 GTCAGTGGGCAAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4080
 Db 2074 GTCAGTGGGCAAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2133
 Qy 4081 GTCATTACGACTCTCTGAGTGGGCTCATGGTGTGCTGCTGCTGCTGCTGCTGCTGCT 4140
 Db 2134 GTCATTACGACTCTCTGAGTGGGCTCATGGTGTGCTGCTGCTGCTGCTGCTGCTGCT 2193

Qy 4141 AATGTCAACTCAGGATGCTCTACTTGGCCCTGATCTGGTGTTCATGAGTACGCGATG 4200
 Db 2194 AATGTCAACTCAGGATGCTCTACTTGGCCCTGATCTGGTGTTCATGAGTACGCGATG 2253
 Qy 4201 CACAAGTCCCGGATGACAGCAGTGTGCGAATGAGGACCTCTCTCAAGATTTGGA 4260
 Db 2254 CACAAGTCCCGGATGACAGCAGTGTGCGAATGAGGACCTCTCTCAAGATTTGGA 2313
 Qy 4261 TGGCTCCAAATCACCCTCCAGGAATTCCTGTCATGAAGCACTGCTACTCTTCAAGCAT 4320
 Db 2314 TGGCTCCAAATCACCCTCCAGGAATTCCTGTCATGAAGCACTGCTACTCTTCAAGCAT 2373
 Qy 4321 ATTCCAGTGGTGGCTGAAATCAAAATTTTGTGATGAATTCGAAATGAATACATC 4380
 Db 2374 ATTCCAGTGGTGGCTGAAATCAAAATTTTGTGATGAATTCGAAATGAATACATC 2433
 Qy 4381 AAGGAACCTGATCGATGATCAATTCGATGCAAAAGAAATCCACATCTCTGCTCAAGAGC 4440
 Db 2434 AAGGAACCTGATCGATGATCAATTCGATGCAAAAGAAATCCACATCTCTGCTCAAGAGC 2493
 Qy 4441 TTCTACAGCTCACCAGCTCTCTGGACTCCGTCAGCTTATTCGAGAGAGCTCATCAG 4500
 Db 2494 TTCTACAGCTCACCAGCTCTCTGGACTCCGTCAGCTTATTCGAGAGAGCTCATCAG 2553
 Qy 4501 TTCACTTTGACCTGCTAATCAAGTCAACATGTCAGCTGAGCTTCCGGAATGATG 4560
 Db 2554 TTCACTTTGACCTGCTAATCAAGTCAACATGTCAGCTGAGCTTCCGGAATGATG 2613
 Qy 4561 GCAGAGATCATCTCTGTCAGAGTCCCAAGATCCTTCTGGGAAAGTCAAGCCCATCAT 4620
 Db 2614 GCAGAGATCATCTCTGTCAGAGTCCCAAGATCCTTCTGGGAAAGTCAAGCCCATCAT 2673
 Qy 4621 TTCCACACCCAGTGAAGCATTTGGAAATCCCTATTTCTCACCCAGCTCATGCCCTTT 4679
 Db 2674 TTCCACACCCAGTGAAGCATTTGGAAATCCCTATTTCTCACCCAGCTCATGCCCTTT 2733
 Qy 4680 CAGATGCTCTTCTGCTGTTATTAACCTGCACTACTCTCTGCACTGCTTGGGAAATTC 4739
 Db 2734 CAGATGCTCTTCTGCTGTTATTAACCTGCACTACTCTCTGCACTGCTTGGGAAATTC 2793
 Qy 4740 CTCTATTGATGTACAGTCTGTCATGAAC 4767
 Db 2794 CTCTATTGATGTACAGTCTGTCATGAAC 2821

RESULT 15
 PHU94176 2769 bp mRNA linear PRI 16-SEP-1999
 LOCUS Papio hamadryas androgen receptor mRNA, complete cds.
 DEFINITION U94176
 ACCESSION U94176.1 GI:3861474
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Papio hamadryas.
 Papio hamadryas.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 Cercopithecinae; Papio.

REFERENCE 1 (bases 1 to 2769)
 Choong,C.S., Kempainen,J.A. and Wilson,E.M.
 Evolution of the primate androgen receptor: a structural basis for
 disease
 J. Mol. Evol. 47 (3), 334-342 (1998)
 JOURNAL
 MEDLINE 98404153
 PUBMED 9732460
 REFERENCE 2 (bases 1 to 2769)
 Choong,C.S., Kempainen,J.A. and Wilson,E.M.
 Direct Submission
 Submitted (18-MAR-1997) Laboratories for Reproductive Biology,
 University of North Carolina at Chapel Hill, CB 7500, MSRB, Rm 370.
 JOURNAL
 UNC-CH, Chapel Hill, NC 27599, USA
 FEATURES
 Location/Qualifiers
 1. .2769
 source

us-09-497-822c-18.rge

Fri May 9 14:20:38 2003

Db 2737 ATGCTCTTCTGCTGTATATACTCTGCACACCC 2769

Search completed: May 8, 2003, 18:56:01
Job time : 13416 secs

QY 3604 GGAAGCTGCAAGGTCTCTTCAAAAGAGCGCTGAAGGAAACAGAAAGTACCTGTGGCC 3663
Db 1657 GGAAGCTGCAAGGTCTCTTCAAAAGAGCGCTGAAGTAAACAGAGTACCTGTGGCC 1716
QY 3664 AGCAGAAATGATTGCACTATTGATAAATCCGAGGAAATTTGCCATCTTGTGCTTT 3723
Db 1717 AGCAGAAATGATTGCACTATTGATAAATCCGAGGAAATTTGCCATCTTGTGCTTT 1776
QY 3724 CGGAAATGTTATGAACGAGGATGACTCTGGAGGCCCGGAGCTGAAGAACTTGGTAAT 3783
Db 1777 CGGAAATGTTATGAACGAGGATGACTCTGGAGGCCCGGAGCTGAAGAACTTGGTAAT 1836
QY 3784 CTGAAACTACAGGAGGAGAGAGGCTTCCAGCACACACAGCCGCCACTGAGGAGACACC 3843
Db 1837 CTGAAACTACAGGAGGAGAGAGGCTTCCAGCACACACAGCCGCCACTGAGGAGACACC 1896
QY 3844 CAGAACTGACAGTGTACACATTTGAAGCTATGAATGTACGCCATCTTTCTGAATGTC 3903
Db 1897 CAGAACTGACAGTGTACACATTTGAAGCTATGAATGTACGCCATCTTTCTGAATGTC 1956
QY 3904 CTGGAAGCCATTGAGCCAGGTGTAGTGTGTGTSACAGACAAACAGCCGCCACTTCC 3963
Db 1957 CTGGAAGCCATTGAGCCAGGTGTAGTGTGTGTSACAGACAAACAGCCGCCACTTCC 2016
QY 3964 TTTGAGCCTTGTCTCTAGCCCTCAATGAAGTGGAGAGACAGCTTGTACACGTGTC 4023
Db 2017 TTTGAGCCTTGTCTCTAGCCCTCAATGAAGTGGAGAGACAGCTTGTACATGTGTC 2076
QY 4024 AAGTGGGCAAGGCTTGCCTGGCTTCCGCAACTTACACGTGAGACACAGATGGCTGTC 4083
Db 2077 AAGTGGGCAAGGCTTGCCTGGCTTCCGCAACTTACACGTGAGACACAGATGGCTGTC 2136
QY 4084 ATTCAGTACTCTGATGGGCTCATGGTGTTCGCATGGCTGGCATCTTTCACCAAT 4143
Db 2137 ATTCAGTACTCTGATGGGCTCATGGTGTTCGCATGGCTGGCATCTTTCACCAAT 2196
QY 4144 GTCAACTCCAGGATGCTTACTTCCGCCCTGATCTGTTTTCAATGAGTACCGATGCAC 4203
Db 2197 GTCAACTCCAGGATGCTTACTTCCGCCCTGATCTGTTTTCAATGAGTACCGATGCAC 2256
QY 4204 AAGTCCCGGATGACAGCCAGTGTGTCGAAATGAGGACCTCTCTCAAGATTTGGATGG 4263
Db 2257 AAGTCCCGGATGACAGCCAGTGTGTCGAAATGAGGACCTCTCTCAAGATTTGGATGG 2316
QY 4264 CTCCAATACACCCAGGAAATCTCTGTCATGAAAGCACTGCTACTCTTCAGCATTAAT 4323
Db 2317 CTCCAATACACCCAGGAAATCTCTGTCATGAAAGCGCTACTCTTCAGCATTAAT 2376
QY 4324 CCAGTGGATGGGTGAAATCAAAATCTTTGATGAACTCGAATCACTACATCAAG 4383
Db 2377 CCAGTGGATGGGTGAAATCAAAATCTTTGATGAACTCGAATCACTACATCAAG 2436
QY 4384 GAATCGATCGTATCATGTCATCAAAAGAAATCCACATCTCTGTCAGAGCGTTC 4443
Db 2437 GAATCGATCGTATCATGTCATCAAAAGAAATCCACATCTCTGTCAGAGCGTTC 2496
QY 4444 TACCAGCTCACCAGCTCTCGGACTCCGTGAGCCTATTGCGAGAGAGCTGATCAGTTC 4503
Db 2497 TACCAGCTCACCAGCTCTCGGACTCCGTGAGCCTATTGCGAGAGAGCTGATCAGTTC 2556
QY 4504 ACTTTGACCTGCTTAATCAAGTACACATGTTGAGCGTGGACTTTCCGGAAATGATGCA 4563
Db 2557 ACTTTGACCTGCTTAATCAAGTACACATGTTGAGCGTGGACTTTCCGGAAATGATGCA 2616
QY 4564 GAGATCATCTCTGTGCAAGTCCCAAGATCTTTCTGGGAAAGTCAAGCCCATCTATTC 4623
Db 2617 GAGATCATCTCTGTGCAAGTCCCAAGATCTTTCTGGGAAAGTCAAGCCCATCTATTC 2676
QY 4624 CACACCCAGTGAAGCATTTGGAA-CCTATTTCACCCAGCTCATGCCCTTTTCAG 4682
Db 2677 CACACCCAGTGAAGCATTTGGAAATCCCTATTTCCTACCCAGCTCATGCCCTTTTCAG 2736
QY 4683 ATGCTCTTCTGCTGTATATACTCTGCACATC 4715